cop 2

JAN 21 1929

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN
FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XX
No. 1

JANUARY, 1929

$5.00 a year.
70c per copy.

CONTENTS

STUDIES IN CHILD PERSONALITY. I: A STUDY OF THE LANGUAGE OF KINDERGARTEN CHILDREN. *Harold Rugg, Louise Krueger, and Arensa Sondergaard* 1

PRACTICE VS. MOTIVATION. *Percival M. Symonds and Doris Harter Chase* 19

A NEW APPARATUS FOR PLOTTING AND A CHECKING METHOD FOR SOLVING LARGE NUMBERS OF INTERCORRELATIONS. (CONCLUDED.) *L. Dewey Anderson and Herbert A. Toops.* 36

A SIMPLIFIED METHOD OF COMPUTING MULTIPLE CORRELATION CONSTANTS. *Frank S. Salisbury* 44

CONCERNING THE STANDARD ERROR OF A DIFFERENCE. *Helen M. Walker.* 53

GRAPHICAL COMPARISON OF TWO FORMULAS FOR THE STANDARD ERROR OF A DIFFERENCE. *Dorothy C. Rowell* 61

THE DIFFICULTY OF AN ITEM. *Elisha F. Bliss, Jr.* 63

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES. 67

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION. 74

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street — Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

LABORATORY EXERCISES IN EDUCATIONAL STATISTICS

With Tables

by ROBERT LEE MORTON

Professor of Mathematics, College of Education, Ohio University

These exercises, the outgrowth of seven years' teaching experience, provide the practice in elementary statistics needed to supplement classroom lectures, textbook assignments, and class discussions. Tables also bound separately.

Now Ready

SILVER, BURDETT AND COMPANY

New York Newark Boston Chicago San Francisco

The Commonwealth Teacher Training Study

By W. W. CHARTERS *and* DOUGLAS WAPLES

Materials for Developing Teacher-Training Curricula and Methods of Supervision.

To be published in February
$3.00

THE UNIVERSITY *of* CHICAGO PRESS
CHICAGO - - - ILLINOIS

TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books, or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $5.00 a year, payable in advance. Foreign postage is 30 cents extra. Single current issues are 70 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered. Back volumes $5.50 each; single back issues 90 cents each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

VOL. XX 1929

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

BOARD OF EDITORS:

HAROLD RUGG, *Chairman*
Lincoln School of Teachers College
Teachers College, Columbia University

JAMES CARLETON BELL
College of the City of New York
FRANK NUGENT FREEMAN
University of Chicago
ARTHUR IRVING GATES
Teachers College, Columbia University
VIVIAN ALLEN CHARLES HENMON
Yale University
RUDOLF PINTNER
Teachers College, Columbia University
BEARDSLEY RUML
General Education Board
LEWIS MADISON TERMAN
Leland Stanford University
EDWARD LEE THORNDIKE
Teachers College, Columbia University

Published Monthly Except June to August by
WARWICK and YORK, Inc.,
Baltimore, Md.

Education

TABLE OF CONTENTS

VOLUME XX (1929) THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

Are Boys Excelling Girls in Geometric Learning? WINONA M. PERRY 270–279
Birth Order and Intelligence. L. L. THURSTONE and RICHARD L. JENKINS 641–651
Children's Interest in Collecting. M. T. WHITLEY. 249–261
Child's Early Experience with the Letter A, The. ERNEST HORN 161–168
Choice of Items for a Test on the Basis of Difficulty. PERCIVAL M. SYMONDS. 481–493
Clinical Observation of Eye Movements in the Rating of Reading Ability. W. R. MILES and DAVID SEGEL. . . . 520–529
Comment on Professor Spearman's Note. EDWIN B. WILSON 217–223
Comparison of Certain Phases of Musical Ability of Colored and White Public School Pupils, A. C. T. GRAY and C. W. BINGHAM 501–506
Comparison of Two Types of Silent Reading as Used by Children in Different School Grades, A. C. T. GRAY. . 169–176
Concerning the Standard Error of a Difference. HELEN M. WALKER 53–60
Constancy of Imaginal Content, The. HENRY BOWERS . . 295–298
Corrections 560
Computation of Correlation Coefficients. EDWARD E. CURETON. 588–601
Difficulty of an Item, The. ELISHA F. BLISS, JR. 63–66
Effect of the Order of Printed Response on an Interest Questionnaire, The. C. O. MATHEWS 128–134
Environmental and Character Factors Involved in Scholastic Success: 1926–1927 Data. FRANK K. SHUTTLEWORTH. . 424–433
Erroneous First Impressions on Objective Tests. C. O. MATHEWS. 280–286
Evaluation of Self-administering Spelling Tests, The. RUDOLF PINTNER, HENRY D. RINSLAND and JOSEPH ZUBIN 107–111
Examination as an Aid to Learning. ARTHUR T. JERSILD . 602–609
Experiment in Verbal Learning, An. GEORGE D. STODDARD 452–457
Experimental Study of Bright and Dull High School Pupils, An. W. H. PYLE and G. H. SNADDEN 262–269
Factors Influencing Retention of History in the Sixth, Seventh and Eighth Grades. SARAH JANET BASSETT. . . 683–690
First Impression versus Second Thought in True-false Tests. MARY LOUISE LOWE and C. C. CRAWFORD 192–195
Fundamental Considerations in the Study of Musical Ability. H. LOWERY. 584–587
Graphical Comparison of Two Formulas for the Standard Error of a Difference. DOROTHY C. ROWELL 61–62
High School Student and Scientific Method, The. M. LOUISE NICHOLS 196–204

How Effective is Specific Training in Preventing Loss Due to the Summer Vacation? L. D. MORGAN 466–471
Incidental Learning. RAYMOND R. WILLOUGHBY 671–682
Index of Integration, An. JOHN W. DICKEY 652–660
Influence of Affective Tone on Learning and Retention, The. RUTH M. CHANEY and ALVHH R. LAUER 287–290
Influence of Age of Pubescence upon the Physical and Mental Status of Normal School Students, The. MORRIS S. VITELES 360–368
Law of Use, The. NOEL B. CUFF 438–447
Measurement of Achievement in General Inorganic Chemistry, The. VICTOR H. NOLL. 458–465
Mental Growth Curve for the Binet Tests, The. L. L. THURSTONE and LUTON ACKERSON. 569–583
Method for Correcting Coefficients of Correlations for Heterogeneity in the Data, A. MARK A. MAY 417–423
"Mysterious" Tropisms: An Illustration of the Gestalt Law of Precision. MARION E. GRANT 125–127
New Apparatus for Plotting and a Checking Method for Solving Large Numbers of Intercorrelations, A. L. DEWEY ANDERSON and HERBERT A. TOOPS. 36–43
New Publications in Educational Psychology and Related Fields of Education. Conducted by FRANCES M. FOSTER . 74–80, 151–158, 231–240, 316–320, 392–400, 472–476, 552–560, 636–640, 717–720
Nomograph for Estimating the Reliability of a Test in One Range of Talent when Its Reliability Is Known in Another Range, A. EDWARD E. CURETON and JACK W. DUNLAP. 537–538
Note on Professor Freeman's Discussion of the Stanford Study of Foster Children. BARBARA S. BURKS 98–101
Note on Reliability Predictions for the Seashore Music Tests. LYLE H. LANIER. 691–692
Note on the Correctness of Certain Error Formulas, A. HARL R. DOUGLASS 434–437
Note on the Correctness of Certain Error Formulas, A. KARL J. HOLZINGER 669–670
Note on the Standard Error of a Reliability Coefficient Estimated from the Coefficient for a Different Range of Talent. JACK W. DUNLAP and EDWARD E. CURETON. 705–706
Notes on Articles in Educational Psychology in Current Issues of Other Magazines. Reported by JAMES E. MENDENHALL. 67–73, 144–150, 224–230, 310–315, 387–391, 542–551, 630–634, 707–716
Objective Measure of Ability to Make Topical Outlines, An. HELEN FERN DARINGER 112–118
On Formula for Estimating the Reliability of Test Batteries. HARL R. DOUGLASS and FREDERICK W. COZENS. 369–377
On Tetrad Differences with Overlapping Variables. KARL J. HOLZINGER 91–97

On the Determination of Reliability in Comparing the Final Mean-scores of Matched Groups. E. F. LINDQUIST and R. R. FOSTER 102–106
Performance Tests and Foreign Children. OLIVE P. LESTER 303–309
Photographic Measures of Reading Ability. MILES A. TINKER 184–191
Photographic Records of Eye-movements in Studying Spelling. ETHEL M. ABERNETHY 695–701
Practice vs. Motivation. PERCIVAL M. SYMONDS and DORIS HARTER CHASE 19–35
Predicting Ability in Art. ALFRED S. LAWERENZ 702–704
Prognostic Value of the Low Range Army Alpha Scores, The. ALFRED H. JONES. 539–541
Publications Received 158–160, 477–480
Relation of Left-Handedness to Psycho-neurotic Traits, The. R. G. WETMORE and G. H. ESTABROOKS 628–629
Relationship between IQ and Size of Family, The. H. E. G. SUTHERLAND 81–90
Relationship between the Techniques of Partial Correlation and Path Coefficients, The. E. LOWELL KELLY. 119–124
Relationship of Motor Ability to the Intelligence of Kindergarten Children, The. OSCAR E. HERTZBERG. 507–519
Relative Effect of Nature and Nurture Influences on Twin Differences, The. KARL J. HOLZINGER. 241–248
Reliability of the Goodenough Drawing Test with Feeble-minded Subjects, The. LLOYD N. YEPSEN 448–451
Reply to Dr. Lanier's Note on the Seashore Tests. PAUL R. FARNSWORTH 693–694
Response to T. Kelley. C. SPEARMAN 561–568
Results of Rescoring Five Hundred Thirty Dearborn Tests, The. WALTER F. DEARBORN and C. WILSON SMITH. . . 177–183
Short Method for Computing the Correlation between Interchangeable Variables, A. FLORENCE L. GOODENOUGH . . 386
"Sight Reading" Method of Teaching Reading, as a Source of Reading Disability, The. SAMUEL T. ORTON. 135–143
Simplified Method of Computing Multiple Correlation Constants, A. FRANK S. SALISBURY. 44–52
Some Conclusions from a Comparison of School Achievement of Certain Racial Groups. HELEN G. PRATT 661–668
Some Results of the Multi-mental Test. PAUL A. WITTY and J. F. TAYLOR 299–302
Some Scientific Approaches toward Vocational Guidance. CHARLES E. LIMP 530–536
Special Case ofthe Multiple Correlation Coefficient, A. A. K KURTZ 378–385
Spelling by Rule. CORNELIA R. TROWBRIDGE. 205–211
Studies in Child Personality. I: A Study of the Language of Kindergarten Children. HAROLD RUGG, LOUISE KRUEGER and ARENSA SONDERGAARD 1–18

Study of the Effect of a Definite Course of Reading in General Literature upon Achievement in Content Subjects with Children of Superior Mental Ability, A. CORA LEE DANIELSON 610–621
Study of the Mental Growth Curve with Special Reference to the Results of Group Intelligence Tests, A. CHARLES LEONARD ODOM 401–416
Subject-matter Progress of Three Activity Schools in Hawaii, with a Note on Statistical Technique, The. HELEN G. PRATT, JACK W. DUNLAP and EDWARD E. CURETON . . . 494–500
Systematic Differences in the Various Parts of the Herring Revision of the Binet-Simon Test when Applied to Normal Dull Adults. H. H. REMMERS 622–627
Transfer of Training in the Cancellation Experiment. B. H. NEWKIRK and R. H. GUNDLACH 291–294
Uniqueness of "G," The. C. SPEARMAN 212–216
Use of Questions in Social Science Material, The. JOHN N. WASHBURNE. 321–359

BOOK REVIEWS

Brownell, William A., *The Development of Children's Number Ideas in the Primary Grades.* (Winona M. Perry). 154
Burton, W. H., *The Nature and Direction of Learning.* (Peter Sandiford) 554
Chadwick, Mary, *Difficulties in Child Development.* (H. Meltzer) 555
Clapp, F. L., Chase, W. L. and Merriman, Curtis, *Introduction to Education.* (Laurance F. Shaffer). 395
Copeland, Edwin Bingham, *Natural Conduct.* (H. Meltzer). . 77
Crawford, Albert B., *Incentives to Study: A Survey of Student Opinion.* (Ben D. Wood) 476
Dearborn, Walter F., *Intelligence Tests. Their Significance for School and Society.* (Gertrude H. Hildreth). 318
Decroly, O. and Buyse, R., *La Pratique des Tests Mentaux.* (R. Pintner) 234
Descoeudres, Alice, *The Education of Mentally Defective Children.* (Translated from the Second French Edition by Ernest F. Row.) (Gertrude Hildreth) 557
Edwards, A. S., *Fundamental Principles of Learning and Study.* (Revised Edition.) (C. C. Crawford) 320
Fletcher, John Madison, *The Problem of Stuttering.* (Elizabeth D. McDowell) 156
Furfey, Paul Hanly, *Social Problems of Childhood.* (H. Meltzer) 555
Good, Carter V., *Teaching in College and University.* (C. C. Crawford). 719
Hart, Joseph K., *A Social Interpretation of Education.* (Harvey Zorbaugh) 474

Hartshorne, Hugh and May, Mark A., *Studies in Deceit.* (Frank K. Shuttleworth) 151
Hollingworth, Leta S., *The Psychology of the Adolescent.* (Laurance F. Shaffer) 78
Horton, Ralph E. *Measurable Outcomes of Individual Laboratory Work in High School Chemistry.* (J. H. Coleman) 157
Hugon, Paul D., *Our Minds and Our Motives.* (C. S. Slocombe) . 155
Hull, Clark L., *Aptitude Testing.* (E. R. Guthrie) 236
Hunter, Walter S., *Human Behavior.* (Ralph B. Spence) 153
Kandel, I. L., *Educational Yearbook of the International Institute of Teachers College, Columbia University.* (Peter Sandiford) . . . 316
Kelley, T. L., *Scientific Method: Its Function in Research and Education.* (Peter Sandiford) 552
Klapper, Paul, *Contemporary Education; Its Principles and Practice.* (I. L. Kandel) 636
Knight, Frederic B. and Behrens, Minnie S., *The Learning of the One Hundred Addition Combinations and the One Hundred Subtraction Combinations.* (W. J. Osburn) 398
Koos, Leonard V., *The Questionnaire in Education.* (C. O. Mathews) 559
Levine, Albert J. and Marks, Louis, *Testing Intelligence and Achievement.* (Gertrude Hildreth) 638
Miller, Nathan. *The Child in Primitive Society.* (Laurance F. Shaffer) 718
Monroe, Walter S. and Weber, Oscar F., *The High School.* (J. H. Coleman) 79
Murphy, Gardner, *An Historical Introduction to Modern Psychology.* (With a Supplement by Heinrich Kluver) (C. S. Slocombe) 473
Orata, Pedro Temesis, *The Theory of Identical Elements.* (Herbert A. Carroll) 636
Osburn, W. J., *Overlappings and Omissions in Our Courses of Study.* (Henry Harap) 317
Pillsbury, Walter B. and Meader, Clarence L., *The Psychology of Language.* (Lester Wilson) 75
Pintner, R., *Educational Psychology: An Introductory Text.* (Peter Sandiford) 472
Purdom, T. Luther, *The Value of Homogeneous Grouping.* (Edward A. Lincoln) 392
Rexroad, Carl Newton, *General Psychology for College Students.* (Melvin Rigg) 637
Rugg, Harold and Shumaker, Ann, *The Child-centered School.* (George S. Counts) 233
Sandiford, Peter, *Educational Psychology, and Objective Study.* (A. I. Gates) 231
Sherman, Mandell and Sherman, Irene Case. *The Process of Human Behavior.* (H. Meltzer) 717
Symonds, Percival M., *The Nature of Conduct.* (Goodwin B. Watson) 74

Thomson, M. K., *Springs of Human Action.* (C. S. Slocombe). . 399
Troland, Leonard T., *The Fundamentals of Human Motivation.* (C. S. Slocombe). 394
Washburne, Carlton and Stearns, Myron M., *Better Schools.* (J. H. Coleman) 396
Youtz, Philip N., Edited by, *An Outline of Aesthetics.* Edman, Irwin, Vol. 1: *The World, the Arts and the Artists.* Wells, Henry, Vol. 2: *The Judgment of Literature.* Allen M. Cecil, Vol. 3: *The Mirror of the Passing World.* Ladd, Henry, Vol. 4: *With Eyes of the Past.* Munro, Thomas, Vol. 5: *Scientific Method in Aesthetics.* (B. J. R. Stolper). 238

INDEX TO AUTHORS

Abernethy, Ethel M. 695
Ackerson, Luton. 569
Anderson, L. Dewey. 36
Bassett, Sarah Janet. 683
Bingham, C. W. 501
Bliss, Elisha F., Jr. 63
Bowers, Henry. 295
Burks, Barbara, S. 98
Chaney, Ruth M. 287
Chase, Doris Harter. 19
Cozens, Frederick W. 369
Crawford, C. C. 192
Cuff, Noel B. 438
Cureton, Edward E. 494, 537, 588, 705
Danielson, Cora Lee. 610
Daringer, Helen Fern. 112
Dearborn, Walter F. 177
Dickey, John W. 652
Douglass, Harl R. 369, 434
Dunlap, Jack W. 494, 537, 705
Estabrooks, G. H. 628
Farnsworth, Paul R. 693
Foster, R. R. 102
Goodenough, Florence L. 386
Grant, Marion E. 125
Gray, C. T. 169, 501
Gundlach, R. H. 291
Hertzberg, Oscar E. 507
Holzinger, Karl J. 91, 241, 669
Horn, Ernest. 161
Jenkins, Richard L. 641
Jersild, Arthur T. 602
Jones, Alfred H. 539
Kelly, E. Lowell. 119
Krueger, Louise. 1
Kurtz, A. K. 378
Lanier, Lyle H. 691
Lauer, Alvhh R. 287
Lester, Olive P. 303
Lewerenz, Alfred S. 702
Limp, Charles E. 530
Linquist, E. F. 102
Lowe, Mary Louise. 192
Lowery, H. 584
Mathews, C. O. 128, 280
May, Mark A. 417
Miles, W. R. 520
Morgan, L. D. 466
Newkirk, B. H. 291
Nichols, M. Louise. 196
Noll, Victor H. 458
Odom, Charles Leonard. 401
Orton, Samuel T. 135
Perry, Winona M. 270
Pintner, Rudolf. 107
Pratt, Helen G. 494, 661
Pyle, W. H. 262
Remmers, H. H. 622
Rinsland, Henry D. 107
Rowell, Dorothy C. 61
Rugg, Harold. 1
Salisbury, Frank S. 44
Segel, David. 520
Shuttleworth, Frank K. 424
Smith, C. Wilson. 177
Snadden, G. H. 262
Sondergaard, Arensa. 1
Spearman, C. 212, 561
Stoddard, George D. 452
Sutherland, H. E. G. 81
Symonds, Percival M. 19, 481
Taylor, J. F. 299
Thurstone, L. L. 569, 641
Tinker, Miles A. 184
Toops, Herbert A. 36
Trowbridge, Cornelia R. 205
Viteles, Morris S. 360
Walker, Helen M. 53
Washburne, John N. 321
Wetmore, R. G. 628
Whitley, M. T. 249
Willoughby, Raymond R. 671
Wilson, Edwin B. 217
Witty, Paul A. 299
Yepsen, Lloyd N. 448
Zubin, Joseph. 107

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XX | January, 1929 | Number 1

STUDIES IN CHILD PERSONALITY

I

A STUDY OF THE LANGUAGE OF KINDERGARTEN CHILDREN

HAROLD RUGG

Professor of Education, Teachers College, Columbia University

With the Collaboration of

LOUISE KRUEGER

Assistant, Teachers College, Columbia University

AND

ARENSA SONDERGAARD

Director of Kindergarten, St. Agnes' Chapel, New York City

I

This is the first of several articles which will report studies of child personality. In it are presented the problems, the psychological theory from which they emerged, and the chief techniques employed. It also reports the first of the detailed studies themselves: A preliminary analysis of the predominant traits of kindergarten children as revealed by their language. Our research program recognizes the imperative need of studies of other important phases of growth—for example, the rôle of the propulsive motives and the orienting attitudes, the growth of meaning, the development of judgment, and the emergence of creative abilities. The studies to be reported this year will be confined to accounts of work done with kindergarten children primarily of the ages of four and five years.

The whole study of the whole child in action—this is the general objective of the investigation.

Child personality can be studied adequately only through the total situation in which it is an emerging nucleus. The life of the kindergarten child is a succession of onward-moving experiences, some fleeting, others sustained. The day is a kaleidescope of evanescent social experiences, each one infinitesimal but crucial; each making its distinctive contribution to the development of individuality. Each of these experiences is a constantly changing complex of motives, attitudes, emotions, physical experiences, meanings, childlike judgments. The contribution that each one makes to personality is determined by the pervasive physical and emotional tone of the childish organism. More specifically it is determined by the integrated action of several factors—*motives or desires*, some obvious, some hidden; motor-mental sets, that is, *attitudes;* and the predominantly potent *meanings* which the child selects and builds into his experience through his reactions.

This hypothesis of the integrative nature of experience is not unique to our study. It is supported by a voluminous body of evidence that has accumulated in the researches of the past twenty-five years. Three independent groups of investigators have adduced an impressive assembly of knowledge: the physiologists and neurologists; the various schools of psychology (the Gestalt and the psychoanalytic, as well as the behaviorist and the scientific schools) and the creative artists and artist teachers. In this series of studies, therefore, we have taken our investigational cues from the integrative nature of experience. Our techniques have been built up on the theory that no method of investigation will be adequate, no matter how objective and scientific it is, which does not study the total conduct of the child in its three-dimensional setting.

It is important, furthermore, to attack the study of child personality from this standpoint because of the great emphasis that has been laid in the past upon physical habits. The literature of the kindergarten abounds with "habit inventories," "room equipment," "environmental set-up," "selection and use of materials." It is conspicuously sparse in the vocabulary of motive, of meaning, of judgment, and of the directive and orienting attitudes. We urge upon those who are bringing about the reconstruction of the kindergarten a more balanced view of psychology, curriculum, and administration. During the past twenty-five years, indeed, elementary education has been swayed by the analytical, statistical methods of the stimulus-response psychology. This has tended to minimize

the significance of the integrative nature of personality. It has tended to emphasize skill, habit, and to veil the rôle of motive, attitude, meaning, and generalization. Our psychology, indeed, must be an eclectic one. It must be constructed from the contributions of Freud, Jung, Adler, and their colleagues in the field of motive; from those of the Gestalt psychologists with their emphasis upon the rôle of relationship in meaning and reasoning, as well as from the offerings of the behaviorists and scientific psychologists with their stress upon the objective study of stimulus and response.

II

The Technique of the Eye-witness Analyst

To make as complete a verbal picture as possible of the succession of experiences through which the child's personality develops, we have sought to employ an inclusive, an eclectic technique, one which would produce a verbatim account of everything the child says and does. For convenience we shall refer to this as the technique of the eye-witness analyst.

To study objectively the whole child in action we need to provide in the classroom two trained and experienced persons—a teacher and an analyst, partners in the completest sense—each making her contribution to the construction of the activity curriculum of the school and each contributing to the analysis of the personalities of the individual children.

The technique employed in these studies is really a synthesis of methods of investigation which have been developed during the past fifteen years. Rugg and Clark devised the method of the eye-witness analyst in their original study of ninth grade mathematics in 1916.[1]

They used it with Schorling in the investigation of junior high school mathematics in the Lincoln School of Teachers College in 1921. The technique was employed from time to time by the Social Science Research Group in the Lincoln School from 1921 to 1924. The verbatim recording and analysis of children's conversations was employed successfully by Piaget in the *Institut* in Geneva in 1921–1922.[2]

[1] Rugg, Harold and Clark, John R.: "Scientific Method in the Reconstruction of Ninth Grade Mathematics." Department of Education, University of Chicago, 1918.

[2] Piaget, Jean: "The Language and Thought of the Child." Harcourt, Brace and Co. "Judgment and Reasoning in the Child." Harcourt, Brace and Co., New York City.

In the light of the knowledge that accrued from these previous uses of the technique, we have sought to make as complete an eye-witness account as possible of the social situations in the midst of which personalities were developing.

The kindergarten child reveals his traits in three ways—through his speech, through his bodily movements and gestures, and through the products of his creative activity. We have recorded and analyzed, therefore, what the child says, what he does, what he produces. Each of these provides an important avenue for the analysis of child motives, attitudes, meanings, reasonings. In order to make the study significant, what the children said was analyzed always in the light of the entire social situation. Repeatedly it was found that the conversation of the children was significant and could be interpreted only by those who recorded it and only in the reconstruction of the setting in which it was spoken. Herein lies, indeed, one of the chief weaknesses of the many stenographic accounts that have been published of children's conversations. They did not include an adequate analysis of the social background in which the conversation took place.[1]

Our experience in the Chicago and Lincoln School experiments proves conclusively that a teacher cannot teach and analyze learning at the same time. The tasks of the teachers on the one hand and of the analyst on the other are distinctly different. The teacher is constantly on the alert to discover interests and needs, to seize upon these and to direct them in the most effective manner into more mature and expanding activity. Her mind is focused upon growth, the re-direction of child activity. She confronts the need of thinking

[1] In developing our technique we evaluated the methods employed in a number of the progressive schools and in the preparing of "diary records" in certain kindergartens. Neither of these contributed procedures of scientific value. The former methods did not give eye-witness accounts of the children in action; rather the accounts were brief wishful descriptions of what the teacher remembered from the gross and more prominent situations of the day, week, or month which preceded. Similarly, the techniques employed in the making of diary records produced little of scientific importance. The diary records, indeed, have been obtained in the pursuit of the worthy aim of the directions of kindergarten-primary departments to train students of education in more careful methods of observing children. They were made by untrained students varying widely in teaching experience, maturity and in grasp of psychology and scientific method; the record was non-verbatim, partial, fleeting; the observers had not cooperated in setting up the activities of the curriculum, hence failed to comprehend the conduct of the child in the light of the total social situations; finally, observations were made in isolated intervals.

and feeling as the child does, as well as of informing him and of drawing him out.

The eye-witness analyst on the other hand is primarily a student of the psychology of learning. Furthermore, released from the necessities of teaching she is in a strategic position to discover how the child learns. She records with expert stenographic assistance, everything that a particular group of children says and does. (This is necessary in order to obtain a verbatim record of what takes place in the succession of social experiences that comprise the day of the kindergarten.) First, then, the analyst is an eye-witness recorder; second, she is an expert interpreter of the child's conduct in action; she is concerned with discovering how the child learns or why he does not. It is clear, therefore, that her purposes and functions are to be clearly distinguished from those of the teacher. We reiterate a teacher cannot teach and record and analyze child experiences at the same time.

An important question inevitably arises in the mind of the reader at this point: Can such a complete record of the educative situation be made without seriously disturbing it? At the end of three months of recording, it has been shown clearly that the records have been made almost without the children being aware of the process. Even the dictation of the analyst to the stenographer was accepted by the children as so much conversation between the adults and as foreign to the children's interest.

III

The Present Study: An Analysis of the Language of Kindergarten Children

Our initial use of the eye-witness technique eventuated in a study of the traits of kindergarten children as revealed through their conversation and the situations in which it took place. The language of the child is undoubtedly the most important single overt exponent of his personality. The kindergarten child is engaged in experimenting verbally with his environment. He is attaching meanings to the persons, objects and actions that compose it. He is a great linguistic experimenter. He revels in the rhythmic grouping of sounds. He toys with novel meanings. He delights in rhythm and rhyme. He is in the mulling-over stage of vocabulary development. Through the study of his language, therefore, we make an effective approach to the comprehension of child nature.

How the Conversation of the Children Was Recorded

The conversation of each of the twenty-seven children was recorded in three fifteen-minute periods: (1) In outdoor play; (2) in word period; (3) in luncheon discussion. In the outdoor play, the children are engaged in a great variety of undirected activities. Each child is left free to choose his own play, engaging either in physical activity about the yard and on the apparatus or in using the apparatus to further his dramatic play. Similarly, in the indoor work period it is the spontaneous and personal interests of the children that dictate what is said and done. Here, too, they are left free to do constructively whatever they are interested in and to choose to work with any materials that are available in the room. Hence the relatively undirected nature of the work period.

Conversation in the luncheon discussion, however, was "directed" in a limited way—to the extent that the mere presence of an adult as leader tends at once to inhibit certain types of conversation and to draw out others. It should be noted in this connection that another phase of the investigation explored children's conversations when they were controlled. This was done in that period of the kindergarten day known as "group discussion." The data of that part of the investigation will be reported in another article.

The conversations were recorded by having the analyst and the stenographer follow a particular child through a series of fifteen-minute intervals on different days scattered over a period of eleven weeks (September 27 to December 14, 1928). In those intervals everything that he said and everything said in his immediate neighborhood was recorded. In addition, elaborate notes were made concerning the actions of the children—a description of the background against which the conversation went on.

During these eleven weeks most of the children were living their first experiences in school. The curriculum of the St. Agnes Kindergarten consists of a steadily developing and maturing stream of individual and group activities. These activities emerge apparently spontaneously from an environment which has been subtly planned by the teacher in the light of her previous experiences with kindergarten children. This environment has been carefully built up to include toys, play apparatus of various kinds, building blocks, dolls and doll housekeeping equipment, clay, crayons, paints, wood and simple wood-working tools, a work bench, sand, books, pictures and other paraphernalia suited to the variety of needs and interests of

young children. The activities developed, therefore, through the response of the individual children to the consistently planned and frequently changed objective environment around them.

Through the interaction of the children with this interesting and wide range of materials developed a succession of individual and group activities, some short-lived, some more permanent, some recurring in nature. Through three stages these activities appear to pass. First, tentative experimenting with materials, getting acquainted with them, as in handling blocks, patting clay, coloring with crayons, paints, etc.; second, attaching meanings to the things they have made; third, the use of materials to further dramatic play, as in the making of a boat, playing house, store and the like. Suffice it to say that the curriculum was free from coercion and composed of relatively individualistic, spontaneous activities directed subtly by the teacher without the children being conscious of the plan. It was in a series of spontaneous situations, therefore, that the conversation of these children was recorded.

Additional types of information were collected concerning the children in order to throw further light upon the development of their personalities. The Stanford-Binet test was given individually to each child. In addition two group tests were taken by the children—the Pintner-Cunningham Primary Mental Test and the Goodenough Drawing Test. The chronological ages, mental ages, and the IQ's will be reported in the next article.

The data show that the children with whom this study was conducted are representative of millions of the children in the public schools of America. The chronological ages range from four years no months to six years four months with a median of five years one month. The IQ's range from seventy-six to one hundred thirty-five, with a median of one hundred eight. The mental ages range from three years ten months to six years five months, median five years seven months. Two of them are children of very well-to-do parents, twelve are from moderately comfortable homes (physicians, dentists, lawyers, etc.), fourteen are from the general skilled artisan-clerical class.

IV

The Scope of the Children's Conversation

Three thousand one hundred twenty-five remarks were spoken by the twenty-seven children in the course of a total elapsed time of

20¼ hours. Each of these remarks was recorded and later analyzed and evaluated as the mark of a given trait. The total of 3125 remarks was composed of two separate samples: first, a sample of 1895 remarks, the total obtained during the time the continuous conversation of each of the twenty-seven children was being recorded; second, a sample of 1230 remarks, the aggregate obtained by recording each of the remarks made by all of the other children in the successive social groups.

The scoring and classification of these two separate sets of remarks enable us to answer the important statistical question of "sampling." Do 1895 remarks, or the grand total of 3125 remarks, constitute a large enough sample to give an adequate measure of the traits of the kindergarten child? At the end of this article Table II presents a comparison of this two-fold classification of the conversations. It shows clearly that the distribution of traits of the first sample of 1895 remarks is approximately the same as that of the second sample of 1230, or of that of the grand total of 3125 remarks. For example, the analysis of the first sample, 1895 remarks, revealed that 43.7 per cent were exhibits of the "self-assertive" traits; the second sample revealed 36.4 per cent. The same close approximation is shown for "linguistic experimentation," 7.3 per cent compared to 4.08 per cent. For "questions asked" 8.7 per cent against 11.8 per cent. For "social consciousness" 3.5 per cent against 4.10 per cent. For "verbalized perceptions" 8.8 per cent against 7.5 per cent. For "dramatic play and the world of fancy" 4.5 against 5.16. It is our conclusion, therefore, that certainly in the total of 3125 remarks we have an adequate "sample" from which to deduce the traits of the kindergarten child.

V

The Traits of the Kindergarten Child So Far as They Are Revealed by His Conversation

We are assuming that the conversation of the children can be evaluated to reveal the traits behind them. Hence each of the 3125 remarks was scored as an index of one or more groups of traits; for example, "self-assertion," "tendencies toward linguistic experimentation," "recognition of inability to achieve," "tendencies toward social consciousness," etc. This extensive evaluation of the remarks of the children resulted, therefore, in an important basic list of traits. This is presented in Table I. The traits have been arranged in the

TABLE I.—CLASSIFICATION OF CHILDREN'S REMARKS

	Number of remarks	Total	Percentage of remarks
I. Self-assertion	...	1275	40.8
1. Expression of personal power	212	....	6.8
(*a*) Prior to achievement	114	....	3.6
(*b*) During achievement	50	....	1.6
(*c*) After achievement	48	....	1.6
2. Display	108	....	3.5
(*a*) Of personal effects	85	....	2.7
(*b*) To exceed personal effects and abilities of others	23	....	.7
3. Interjection of self into situation	133	....	4.3
4. Defense of one's feeling of ownership, resistance to interference	67	....	2.1
5. Commands, giving permission to others	451	....	14.4
6. Threats	56	....	1.8
7. Derision	93	....	2.98
8. Requests	57	....	1.8
9. Contradiction	98	....	3.1
II. Self-depreciation (recognition of inability to achieve)	...	8	.25
III. Evidences of social consciousness	...	116	3.7
1. Enthusiasm for group achievement	2	....	.07
2. Sympathy or concern for others	18	....	.58
3. Spontaneous sharing of experiences or possessions	65	....	2.1
4. Critical evaluation and commendation	23	....	.7
5. Acquired social forms	8	....	.25
IV. Verbalized perceptions conditioned by	...	259	8.3
1. A situation	64	....	2.0
2. An act	16	....	.55
3. A felt need	30	....	.96
4. Physical sensations of sight, touch, smell, hearing	26	....	.83
5. Recognition of size, color, form, number	68	....	2.2
6. Recognition of time and space	5	....	.16
7. Soliloquy accompanying action	50	....	1.6
V. Linguistic experimentation	...	188	6.0
1. Rhythmic grouping of sound	33	....	1.1
2. Play upon a word	77	....	2.5
3. Repetition of sounds heard	5	....	.15
4. Repetition of words heard	31	....	.99
5. Play with rhyme	18	....	.58
6. Recognition of differences in pronunciation	9	....	.2
7. Play upon an idea	15	....	.48
VI. Dramatic play in the world of fancy	...	148	4.7
VII. Questions	...	311	9.95
VIII. Rational thought	...	193	6.2
1. Explanation of procedures or processes	83	....	2.65
2. Justification of thought and conduct	67	....	2.18
3. Contradiction	43	....	1.37
(*a*) With implied causal relationship			
(*b*) With implied logical relationship			
IX. Statements of fact	...	500	16.0
1. Information; description	386	....	13.32
2. Personal experiences	42	....	1.34
3. Conduct	72	....	1.34
X. Answers of "yes" and "no"	...	127	4.1
Total	...	3125	100.0

order in which they are discussed in this article. It is important to note that the analysis of the conversation was made without preconception of the traits producing the remarks. The attempt was made to detect independently for each remark the trait of which it was a measure.

From this list of traits with their indicated frequencies we can draw important inferences concerning the personalities of kindergarten children. The classification provides a tentative mass view of child personality. In a subsequent article the data will be reinterpreted in the attempt to evaluate individual personalities. We shall then be able to discriminate whether a child is assertive, domineering, or submissive, docile. We can detect whether he commands, leads, dominates, or whether he acquiesces, follows, submits. He will be revealed as living in a world of fancy or perhaps merely in the concrete manipulation of things and people.

The Kindergarten Child Essentially a Protagonist of Self

Our data reveal clearly the dominance of the "self-assertive" traits. Two-fifths of the conversation of the children revealed naive interest in themselves and their own affairs (see Table I). Less than one twenty-fifth of the conversation is an overt expression of the child's interest in the group. Not only do the summarized statistics of our study prove this, the very classification itself sets it forth in startling fashion. Of the 3125 remarks, 1275, were descriptive of the traits we have termed "self-assertion"—*e.g.*, of personal power, of self display, of interjection of self into a situation, of defense of one's feeling of ownership, of resistance to interference, of contradiction, of commands, threats, and derision. The photographing of the whole kindergarten child in action reveals him as essentially a defender of his individuality, a non-conformist, a relatively unsocial being.

This fact is revealed in another way in our statistical and graphic exhibits. Only eight remarks in 3125 indicated that the children recognized their own inability to achieve (see Table I).

VI

Two Possible Interpretations of the Facts Concerning the Self-assertive Traits of the Kindergarten Child

The facts which have emerged from the preliminary study are of tremendous significance for the elementary school but their inter-

pretation presents an important controversial issue. They can be interpreted in two opposed ways. One group of kindergarten-primary workers will interpret the dominance of the traits of self-assertion as dangerous to society and will organize its curriculum with the chief goal in mind of making the children socially conscious. These workers regard the primary objective of education as adjustment to society; they seek chiefly to preserve the social heritage; their sole guiding aim is social efficiency, and their oft-reiterated admonition to teachers is to fit the child to conform to the social group.

There is another point of view—the one, indeed, to which we commit ourselves in the development of a curriculum for the elementary school. This point of view regards the appearance of these self-assertive traits with great equanimity. These traits present, indeed, great potentiality for the development of individuality. This school of thought seizes eagerly upon the fact that when children enter school they reveal that the social group has not yet succeeded in stamping in a marked sense of inferiority.

This point of view recognizes, furthermore, that in times past a regime existed and is even widespread today, in which home, neighborhood, the community and the school cooperate to produce, between the ages of five to fifteen, thorough-going conformists. Under the conception of adjustment to the social group the years of elementary education warp and mould individuality to fit conventional mores, beliefs, opinions, norms of conduct. The naive sense of personal individuality with which the child enters the school, therefore, is replaced increasingly by the sense of inferiority and the need for conformity.

The educational theory upon which we shall interpret the data of the traits of the kindergarten child comprehends, however, two great foci: On the one hand, adjustment to society, on the other, the preservation of individuality. The kindergarten must be broadly oriented by the vision of both of these goals. The life of the kindergarten-primary child is essentially individualistic; but it should also be a time in which the demands of social adjustment will be permitted to press in upon him gradually and subtly. It must necessarily be a period in which he is also becoming acquainted with the verbal side of his environment, a stage of linguistic experimentation and of dramatic play.

Types of Conversation Which Reveal the Self-assertive Traits

To some students of the kindergarten our emphasis upon the self-centered nature of the child will appear exaggerated. It is important, therefore, that the reader have at hand typical illustrations of our interpretations of the children's conversation. The first main division of Table I reveals nine categories of conversation all of which appear to us to be examples of the self-assertive traits.

Fourteen per cent of all the remarks of the children were commands. The eye-witness of the social life of these children listens constantly to the reiteration of commands like these: "Get offa there" . . . "Fasten it up, Donald" . . . "Don't push me" . . . "Come here" . . . "Say, hurry up" . . . "Get me offa this ship" . . . "Get in there and put the coal there" . . . "You be in the smoke stack" . . . "You put those balls down" . . . "Come on up in your place" . . .

Interspersed with the commands are frequent expressions of personal power. 6.8 per cent of all the remarks of the children asserted the ability of the individual to achieve or exhibited exultation in possession or achievement. For example: "I had a hard time teaching myself one day, but I learned hard because I learned quick" . . . "I built that one myself, don't you remember?" . . . "I'm head man, and, Dicky, you are one of my soldiers" . . . "Charles is the boss, but I am a little bit bossy, too" . . . "I'm the captain" . . . "I'm the boss airplane" . . . "Look what I did" . . . "Look, I'll show you" . . . "I'm higher than you" . . . "I can climb up another high one" . . .

Most of these children also reveal frequent attempts to interject themselves into group situations. Four per cent of the 3125 remarks were made under such circumstances. These are typical. Each was made in the attempt of a child to include himself in the center of an already established group: "Hey, I'll fix that" . . . "Can I help too?" . . . "I can play with that" . . . Hey, you" . . . "Listen, listen" . . . "Let me see" . . . The reader should be cautioned again that these remarks can be interpreted adequately only by an eye-witness who is alert to the total situation in which the conversation is going on.

The self-assertive nature of the child reveals itself also in remarks of self-display and the desire to exceed others either in achievement

or in personal effects. For example: "Mine knows more than yours" . . . "Wait until you see how nice mine's going to be" . . . "Oh, I'm higher than anybody" . . . "Look at mine" . . .

Individualism appears likewise in resistance to outside interference, in defense of proprietary rights. For example: "This is my track" . . . "Dicky, I was in here first" . . . "Nobody on the boat" . . . "No girls on this boat" . . .

Primitive argument consisting of denial and reiteration was also regarded as an index of the self-assertive nature of the child. Frequently there appears in the record: "It was so" . . . "It was not" . . . "Yes, he can" . . . "Yes, it is" . . . "It is not" . . . "Yes, she does" . . .

Finally, in the conversation of about one-fourth of the class there were many examples of threats and derision: "I'll make you cry like I did ——" . . . "I'm going to pound you" . . . "See what will happen to you tonight" . . . Scoffingly they derided each other: "Old rough Martin can't do that, I'm sure" . . . "But you wouldn't be Jack, would you?" . . . "Oh, you're afraid" . . . "You can't count" . . . "Hey, Baby" . . .

Finally, the second major division of Table I—"Recognition of inability to achieve"—and the third—"Evidences of social consciousness"—illustrate by sharp contrast the conspicuous rôle of self-assertion. In a total of 3125 remarks only eight were clear indications of self-depreciation and these remarks were made by three children. Furthermore, this conclusion is reinforced by the meagerness of evidences of social consciousness revealed by the conversations (see Tables I–II). Only one hundred sixteen remarks were clear indications of expressed enthusiasm for group achievement, sympathy or concern for others, or desire to share experiences or possessions with others. The remarks which we have classified in this category are illustrated by the following: (1) Commendation of others: "——, you certainly found a big napkin all right;" (2) Concern for others: "Do you think I want to make Miss S.'s milk tumble over?" . . . "Did you hurt yourself, honey?" . . (3) Experiences shared with others: "Dicky, do you want to see an airplane on the wall?" . . . "Let's go skipping" . . . "Do you want to play reindeer?" . . . "Let's see if we can find someone else." Finally, occasional remarks like "Excuse me" . . . "Thank you" . . . indicate the beginning of the acquisition of social forms.

The foregoing quotations are thoroughly typical examples of the classification of the remarks.[1]

The Kindergarten Child Employs Speech to Become Acquainted with His Environment

Granted that the child is essentially a self-assertive individual, lacking a sense of inferiority, self-depreciation, and revealing little social consciousness, what other important traits are revealed by his conversation? Two closely related ones are revealed by our data—the employment of speech to become acquainted with his environment and the decided tendency toward linguistic experimentation.

Eight per cent of the conversation of the four- and five-year old children consisted of verbalized perceptions. Children are getting acquainted with their environment by defining it verbally. Note, for example, the varied illustrations in Table I (IV—1 to 7). A child is a member of little group, sitting at a table. Looking about at the other members of the group he comments out loud: "Just eight at this table" (IV—1). While playing in the yard he breaks a stick and says, "I broke it" (IV—2). Another is building with blocks and expresses a felt need with the remark, "It takes a longer block." Still another is drawn to a group of children playing in the sand and says, "I have got to get in the sand" (IV—3).

This is a period of much physical experiencing. Perceptions—visual, auditory, tactual, olfactory—press in upon the children from all sides. One exclaims, "Oh, that hurts!" Another says, "It's a funny way to chew it" (IV—4). Others are recognizing and attaching words to properties of size, color, form, and number. For example, "This cake is a little bit blue" . . . "Two and two is four and eight and eight is nine" . . . "Oh, it's hard" . . . "1, 2, 3, 4" (IV—3 to 5).

The child also brings language to aid his learning processes by commenting upon the movement of time and upon spatial relations. Witness, for example, an interesting study of his recognition of oppo-

---

[1] The reader should not confuse our use of the term "self-assertion" with Professor Piaget's term "ego-centrism." The latter defines ego-centric speech as that in which a child "talks either for himself or for the pleasure of associating anyone who happens to be there with the activity of the moment." Piaget divides all the conversation of children into two large classes; first, the ego-centric, and second, the socialized. In a subsequent article we shall report a new analysis of our data for the purpose of direct comparison with Piaget's results.

sites: "You pointing to me and this one pointing to that and that one pointing to that." Interpreted in the light of the actual situation, this remark meant that the teacher was opposite him, that another child was opposite a third, his neighbor, and so on.

Finally, he is a little soliloquizer, getting acquainted with his environment by talking to it. Throughout the day his comments come, sometimes rarely after long pauses, sometimes in a stream of exclamation: "Let's go, plane" . . . "He makes me go crazy" . . . "Oh, I'm going to get in the water" . . . "I think I'll pull mine way over here" (IV—7).

The Child's Interest in Linguistic Experimentation

Not only is the kindergarten child employing speech to become acquainted with his environment, he is also a great experimenter with the rhythmic manipulation of words. He enjoys rhythmic groups of sound. Witness his frequent play upon words:

"Big cracker Gitz
"Big cracker Gitz
"You, a funny Gitz" (V—1 and 2)

He enjoys repeating sounds heard and words heard:

"Oh, nice baby,
"Oh, nice baby"
"Way, way, way, way, down"
"1, 2, 3, 5, 9" (repetition of numbers heard in any order, but rhythmically)
"Jackie, Jackie" (V—3 and 4)

He plays much with rhyme: "Hity-Tity" (V—5).

He is interested in discriminating differences in pronunciation: "I said 'ämen.'" (All the other children had said "ämen") (V—6).

He enjoys playing upon ideas: "Her mouth is too small for that cup" . . . "I have only got a little bit of water" . . . "Mine is rotten inside" . . . (V—7).

Conversation and the World of Fancy

Closely related to the child's interest in linguistic experimentation —his playing with words, sounds, rhymes, and meaning—is his tendency to live in imaginary situations. Almost five per cent of the conversation of the children consisted in statements of fancy. For example: ——says, talking to his toy, "The mail wagon hasn't come yet" . . . "Hello, mail wagon, are you here yet?" . . . "Oh, how

do you feel, mail box man?" . . . "It's supposed to be a fence on fire" . . . "See this fence all on fire, falling down" . . . "I'm the cop" . . . "Oh, my babies!" . . . "I can't unwrap them" . . . "He can get arrested for that."

Engagements are made in the world of fancy as well as in the world of fact: "At ten o'clock" . . . "Well, come to my house at four o'clock" . . . "Goodbye" . . . "I'll see you tomorrow" . . . "Maybe the mail-box will come tomorrow" . . . (VI).

Is Intellectual Curiosity a Marked Trait of the Kindergarten Child?

Not if his conversation is a sound index of his traits. One-tenth of the remarks of the children consisted of questions (VII). Of the questions asked by the children, less than ten per cent reveal intellectual curiosity. The questions abound with: "Where's your sister?" . . . "Where's Marian?" . . . "What happened to your arm?" . . . "Are we going home?" . . . "Does anybody sit here?" . . . "Are you taking off your glasses?" . . . "Are you passing the milk? . . . "Did you eat yours? . . . "Do you want it open?" . . . Only rarely did questions indicative of a real intellectual curiosity to know "why" occur.

The Faint Emergence of Rational Thought

Closely related to the question of intellectual curiosity is that which deals with the extent to which the child reasons. Although in a later article we shall report at length our detailed studies of reasoning, we have included in Table I the statistical facts which bear upon this question. A special attempt was made to segregate all examples of conversation which clearly revealed reasoning, even of the most limited types. These were brought together in three categories (VIII).

1. Explanation of procedure and processes.
2. Justification of conduct and thought.
3. Contradiction employing either causal or logical relationship.

An important caution, however: These conversations were essentially spontaneous ones. They did not include "group discussions." The group discussion was under the direction of the teacher and revealed a somewhat greater amount of reasoned explanations and a slightly less amount of intellectual curiosity. Conclusions based

upon the data of Table I (VIII) must be taken as tentative and as resting upon an analysis of spontaneous conversations only.

Excluding the remarks classified in (IV)—verbalized perceptions—we conclude that the conversation of children reveals few evidences of reasoning. Approximately six per cent of all the spontaneous conversations of the four- and five-year olds indicated this trait. Of the remarks which could be so classified, one-half consisted of explanations of procedure and processes. For example: "You must put the chair on the platform and you can reach up" . . . L. W. advises a child not to pat the clay so hard . . . "Then you will get it worse" . . . F. X. explains a game . . . "Somebody goes up the stairs like. Then somebody comes down. Then they turn around and say, 'Red Light'" . . . E. E. explains how to lace shoes . . . "Turn it again and cross it again" . . . H. C. explains a game . . . "Whoever gets it wins."

The little individualist reveals in another respect the similarity between his traits and those of his elders; for example, in his frequent justification of his conduct. K. H. on being criticized for not having included windows in a church which he was drawing, defended his

TABLE II.—COMPARISON OF DISTRIBUTION OF REMARKS IN THE FIRST AND SECOND SAMPLE

Classification	Number of remarks, first sample	Per cent	Number of remarks, second sample	Per cent	Totals	Per cent
I	828	43.7	447	36.4	1275	40.8
II	6	.3	2	.16	8	.25
III	66	3.5	50	4.10	116	3.7
IV	168	8.8	91	7.5	259	8.3
V	136	7.3	52	4.08	188	6.0
VI	85	4.5	63	5.16	148	4.7
VII	165	8.7	146	11.8	311	9.95
VIII	122	6.4	71	5.7	193	6.2
IX	239	12.6	261	21.3	500	16.0
X	80	4.2	47	3.8	127	4.1
Totals.....	1895	100.0	1230	100.0	3125	100.0

procedure by the remark: "The windows is all broken and they put stones over the windows so they won't break."[1]

In Conclusion

This preliminary study of children's conversation has revealed certain general outstanding traits. We have found that he is essentially a self-assertive individual; that he is a linguistic experimentalist, engaged in using words to become acquainted with the world around him, that he reveals only slight evidences of intellectual curiosity and little interest or ability in more than the simplest forms of perceptual thinking.

The dominance of these individual traits suggests two procedures for the elementary school. One is to preserve, so far as possible, the developing individuality; the other to bring the individual into a growing social consciousness—that is, to increase the interplay of personality upon personality. We have suggested that neither procedure is complete without the other. In subsequent articles we shall report the manner by which the relatively free activity curriculum conduces to develop the individualistic traits of the child and at the same time to bring him into a more thorough understanding of his environment.

To do this will necessitate the careful study of individual personalities and the use of other approaches than that of language.

---

[1] It will be noted that in the body of our article we have not referred to statements of fact (IX) and answers of "yes" and "no" (X). Five hundred of the 3125 remarks consisted merely of statements of fact, descriptions of personal experiences, information needed by someone in the group. Sixteen per cent of the remarks were of this character. In addition 4.1 per cent were direct answers of "yes" or "no" to questions. It has seemed wisest to exclude these remarks from the classification of traits.

PRACTICE VS. MOTIVATION

PERCIVAL M. SYMONDS AND DORIS HARTER CHASE

Teachers College, Columbia University, N. Y.

One of the distinct cleavages of thought and practice in education concerns the relative efficacy of practice and motivation in learning. There is one school of psychologists which would stress the importance of practice and repetition in learning. For the school the "Law of Use" is of supreme importance. Put into practice in education this eventuates in the method of drill. All other considerations become eclipsed by this importance placed on drill. Arithmetic, handwriting, reading, even geography and hygiene—all school subjects, in fact—become a series of practice exercises. Motivation is not neglected, but it becomes subsidiary—a mere nursemaid to lead the child up to his drill and make him perform it without resentment. Motivation is only a device to make children practice, for by practice alone is learning accomplished.

The other school of thought places the emphasis on motivation. Their argument is that unless practice is accompanied by interest and zeal, indeed, unless practice becomes a *means* for achieving some desired *end*, learning does not result. This school of thought maintains that learning does not take place without motivation and cites as an illustration of their point of view the old story of the boy who was required to remain after school to write "I have gone" one hundred times to teach him to use correct English. After finishing his task, so the story goes, the boy turned his sheet of paper over and wrote "dear teacher, I have went home." The boy in school may fail to know his number combinations after many weary hours of repetition, but learns his football signals in a single afternoon. Increase the motivation and you decrease the necessity of repetition, says this school of thought. This theory is put into effect in education through the project method. The task of education is well on its way when it has created the desire to learn. Practice and drill cannot be eliminated but the amount of drill may be lessened and the quality of learning increased with better motivation.

In order to determine the facts concerning this issue an experiment was undertaken in the field of learning correct English usage. In brief the experiment was planned so that groups of children were subjected to practice in English usage without motivation, and also

with two types of motivation which we call test motivation and intrinsic motivation. Tests were applied before and after the practice and gains were computed. The experiment was so planned that it is possible to measure the effect of different amounts of practice, and the effect of different types of motivation with practice constant. Finally conclusions are drawn concerning the relative effectiveness of practice and of motivation for the learning.

Material

After casting about for suitable material for use in this experiment it was decided to work in the field of English usage. The Charters Diagnostic Language Tests were selected for use. These tests come in four parts: miscellaneous A; miscellaneous B; verbs; pronouns.

The tests consist of forty short sentences, each sentence (except the first) containing an error such as

He is more nobler.
The Indians, why, they ran away.
We was on time.
Him and I will do it.

In taking the tests a pupil has to rewrite the sentences correctly. The tests come in two forms, 1 and 2, exactly similar in content, so that an error in form 1 finds its counterpart in form 2 of the same test.

Example: Pronouns test

Form 1. Item 2.—May Inez and we go?
Form 2. Item 2.—May May and me sweep?

Form 1 was used for testing before and after the practice. Form 2 was rewritten so that all of the sentences contained correct English, and was used for practice material. This practice material was read orally and in unison. Although the tests were written it was thought best not to have the practice in writing for two reasons. First, with unison reading it was possible to insure practical uniformity in the reading and to have guaranteed correct repetition of the practice material. There could have been no corresponding control of the correctness of the copying. Second, reading aloud is more pleasant and less time-consuming than writing and these were two practical considerations both in the conduct and the outcome of the experiment. It is realized that there is certain to be loss in transfer from the spoken practice to the written test, but at least the loss in transfer was

of the same amount in all groups so that the comparisons to be made were not affected.

Since the practice material was read orally, certain items were omitted in which the error was evident by sight alone (such as spelling) and other items in which the error was plainly recognized by sound, as well as by sight, were substituted in their place. Corresponding shifts in the same items were also made in form 1 of the tests.

Place of Experiment

The norms for the Charters Diagnostic Language Tests show that a score of fifty per cent (or twenty items correct) is made by Grade VI children in general. Accordingly, the tests were given in Grade VI in six[1] New York Public Schools as follows:

School	Rooms Used
Manhattan	
P. S. 87	VIB1, VIB2, VIOpp.
P. S. 10	VIB1, VIB3, VIB4
P. S. 165	VIB1, VIB2, VIB3
Brooklyn	
P. S. 114	VIB1, VIA3, VIA4
P. S. 210	VIB1, VIB2, VIB3
P. S. 183	VIA1, VIA2, VIA3

Three sixth grade rooms were used in each school. The regular classroom teachers of each room conducted the testing and carried on the practice. These teachers were given careful instruction in the technique of the experiment and entered heartily into the spirit of it. There is every reason to believe that the detailed directions given to each teacher were carried out correctly.

In the three Manhatten Schools the material was given in the following order of difficulty, the easiest test coming first: miscellaneous B; verbs; pronouns; miscellaneous A. In the Brooklyn Schools the tests were given in exactly the reverse order.

The facts regarding the difficulty of each of the forms and the gain is given in the following table:

---

[1] This study was accomplished only by virtue of the cordial cooperation of the principals of the six schools—Miss Ruth V. Danheiser, Mr. Abraham Ehrenfeld, Mr. Jacob Theobald, Dr. Margaret E. Lacey, Mr. Frederick Graham and Miss Mary A. Kennedy, and by the painstaking work of the eighteen teachers in whose rooms the testing and practice was carried on.

	Test			
	Miscellaneous A (N = 383)	Pronouns (N = 386)	Verbs (N = 382)	Miscellaneous B (N = 378)
Second test	17.441	23.062	25.796	26.577
First test	14.734	21.008	23.908	25.074
Gain	2.707	2.054	1.888	1.503

One type of motivation was used throughout a school, and both types of motivation was used in two schools—one each in Manhattan and in Brooklyn. Since the order of testing was reversed in the two schools in which any type of motivation was used, it was not necessary to correct the gains made for the order in which the tests were given.

The teachers were directed in scoring the tests to disregard all errors of copying or spelling, but to call any item correct in which the incorrect form had been changed to the correct form.

This material, in one way, is not the most suitable for an experiment of this kind. In the first place it is not new material for the children, for naturally the forms and constructions are part of their everyday language and they have had considerable practice in either the correct or incorrect usage. Children taking part in the experiment are at different stages of learning. Again, it would be desirable in such an experiment to select material which would receive practice only as provided by the experiment itself. Naturally these language forms are heard or used outside of school throughout the experimental period. On the other hand it was necessary, on account of the scope of testing required to determine the differences between the methods used, to select subject-matter which is used in school instruction. Whatever validity the findings have depends on the degree to which all groups are equally subjected to outside practice with these forms. The teachers whose classes were used in the experiment were strongly urged to omit the particular constructions used in the tests from the regular English instruction during the experimental period and to be careful not to mention or emphasize these particular usages or errors. With the large numbers used in this experiment there is good reason to believe that no class group was particularly favored or discriminated against because of the use of elements in the tests outside the experimental testing and practice provided.

No effort was made to pair children in classes on the basis of their initial ability. The correlation[1] between initial score and gain is +.138.

The regression equation for predicting gain from initial score is $\bar{Y} = .0474X + 1.158$. A correction was necessary, therefore, to allow for the influence of initial score on gain. The approximate correction to be added to or subtracted from obtained gain, to allow for the influence of initial score, is given in the following table.

Initial Score	Amount to Be Added to or Subtracted from Obtained Gain
30	−.63
25	−.39
20	−.15
15	+.08
10	+.32
5	+.56

An inspection of the results shows that this correction makes the findings more consistent.

Method

The object of the experiment was to determine the influence of practice and of motivation on learning. In brief, the scheme was to test, practice, and repeat the test.

Part I. No Motivation.—In P. S. 87 Manhattan and P. S. 114 Brooklyn, the experiment was conducted without motivation to determine the effect of sheer practice. Naturally the situation had some kind of motivation in it or the children would not have taken the tests or gone through with the practice at all. The work was performed as a regular school exercise and the children were impelled to their task by the personal influence of the teacher and the habitual response of doing what is expected in the school situation. That is, no incentive outside the ordinary school routine was added to induce the pupils to improve on the tests. The directions given to the teachers in the *no-motivation* schools were:

> Follow all directions exactly. During the period in which this testing and practice is being carried on in your class, be careful not to mention or emphasize or answer any question concerning these matters of English usage. It is important

[1] This was found by correlating the seventy-two mean scores of the eighteen classes with the average gains made. This correlation between mean scores and gains equals the correlation between separate scores and gains. See Kelley, T. L.: "Statistical Methods, Formula 118."

that the only practice in these language forms be that provided by the material itself. You can afford to make this concession in your teaching because your pupils will receive more practice in these language forms than you ordinarily give them.

This is an experiment. The object of your work is *not* to see how large a gain your pupils can make on these tests but to see what gain your pupils *will* make using the procedure that is outlined for them. Other procedures are being tried in other schools and we shall want to compare the achievement of this procedure with the achievement made by other procedures. If you do anything except what is defined for you in these directions we shall get a wrong impression as to the value of this procedure.

Do not tell the children what you are planning to do. Do not tell them what they do on any of the tests. Do not tell them that they are going to have any more tests. Do not tell them what the practice exercises are for. Do not tell them that the practice exercises contain the correct forms. If they ask you a question about it tell them to wait and see.

Keep a record of exactly what tests or practice exercises you give, the date and time they are given.

In the *no-motivation* schools the practice on the first test was given *once*, on the second test *thrice*, on the third test *five* times and on the fourth test *ten* times. This increase in the amount of practice was for the purpose of determining the effect of differences in sheer practice with no motivation.

The first test was repeated immediately after it was given the first time in order to determine the practice effect and also to familiarize the children with the taking of the test so that the practice effect would not be a factor in the results. This was done for the first test given in each of the eighteen classes used in the experiment. The practice effect was found to be +.618 for five hundred sixty-nine cases.

The nature of the practice can be described best by giving the directions for conducting the practice as they were given to the teachers.

Distribute the practice sheets—miscellaneous B. Have the class read through the sentences aloud and in unison. See that everyone reads. If the class reads a sentence incorrectly or mispronounces a word have them repeat it (once) correctly. Do not let the children hurry. See that they read slowly and distinctly. A good way to control this is to have the teacher read the number of the sentence before the class reads the sentence. Do not allow the class to repeat any of the sentences. Collect the practice sheets when all of the sentences have been read.

The program for the *no-motivation* classes in P. S. 87 Manhattan was as follows:

Preliminary Test—miscellaneous B	First day—morning.
Initial test—miscellaneous B	(Given immediately after the first test).
Practice sheet—miscellaneous B	First day—afternoon.
Final test—miscellaneous B	Second day—morning.
Initial test—verbs	Third day—morning.
Practice sheet—verbs (three repetitions)	Third day—morning and afternoon. Fourth day—morning.
Final test—verbs	Fourth day—afternoon.
Initial test—pronouns	Fifth day—morning.
Practice sheet—pronouns (five repetitions)	Five consecutive times mornings and afternoons.
Final test—pronouns	Seventh day—afternoon.
Initial test—miscellaneous A	Eighth day—morning.
Practice sheet—miscellaneous A (ten repetitions)	Ten consecutive times mornings and afternoons.
Final test—miscellaneous A	Twelfth day—afternoon.

The order of tests was the opposite in P. S. 114 Brooklyn.

Part II. Test Motivation.—This part of the experiment tried in P. S. 10 Manhattan and P. S. 210 Brooklyn made use of *test motivation.* By test motivation is meant the use of test results to stimulate the pupils to attempt to improve their scores. The pupils were told the scores they made on the initial test and were urged to try to improve score when the test was repeated. They were also given their scores on the final test so that they might know the degree to which they were successful in improving. The directions to the teachers read:

> Tell the children what you are planning to do. Tell them that they will be given some tests to see how correctly they use the English language; that these tests contain wrong language forms which they will be expected to correct; that sheets will be distributed containing the correct forms which they will read aloud and which will give them an opportunity to find how to better their score; and that they will be given a chance to take the test over to see how much better they can do. Do not urge them to use better English or do not discuss with them their use of English. But urge them to try to improve on the tests. The pupils should be told the score that they make on the first tests in each case and also the score that they make on their re-tests.

In P. S. 210 the principal took special interest in the experiment and had graphs constructed which showed the position of each pupil on the test, so that besides the self-competition to improve one's own score there was the added incentive to try to reach a higher place in the class distribution. Class medians were also computed and told to each class so that there was self-competition, individual competition and group competition.

In Part II—the *test motivation* classes—each practice sheet was repeated *three* times for all tests. These results may be compared with the *three* repetition tests in the *no-motivation* classes.

The schedule for the test-motivation classes in P. S. 10 Manhattan was:

1. Give test—miscellaneous B.
2. Repeat test—miscellaneous B.
3. Score test.
4. Give children their score—morning—urge them to try to improve.
5. Use practice sheet—first time—morning—miscellaneous B.
6. Use practice sheet—second time—afternoon—miscellaneous B.
7. Use practice sheet—third time—morning, second day—miscellaneous B.
8. Repeat test—miscellaneous B—morning, third day.
9. Score test.
10. Give children their scores on retest.

Repeat these steps with the omission of step 2 in the same order for *verbs, pronouns, miscellaneous A*.

In P. S. 210 Brooklyn the same schedule was followed but the order of the tests was reversed.

Part III. Intrinsic Motivation.—In the third section of the experiment an attempt was made to interest the children directly in improving their language usage. This is called *intrinsic motivation* although it must be admitted that the actual motivation aroused by the methods employed fell far short of the intrinsic motivation desired. Directions to the teachers in P. S. 165 Manhattan described this part of the experiment as follows:

This is an experiment to determine the value of certain kinds of motivation in learning habits of correct English usage. In your school we shall use "intrinsic" motivation which means that we shall try to interest the pupils in improving their own English. But this motivation must be accomplished without providing any drill on the items of English usage which we are using except that provided by the tests and practice material.

The week before the tests are to be used efforts should be made to stimulate the interest of the pupils in improving their own English. In order to standardize the procedure the following devices should be used:

1. Devote one assembly to a program in which the need for better English is stressed.

2. Let the principal visit each of the classes which are included in the experiment and urge pupils to try to improve in their English usage.

3. Use the sheets entitled "In the Electric Company." In using these sheets let the pupils read the passage through silently. Then have them answer in writing the questions at the end of the passage. Spend from ten to fifteen minutes in class discussing their answers to the questions.

4. Distribute and use the sheets headed "Questions for Discussion" for an impromptu discussion of questions 1 and 2, taking fifteen to twenty minutes.

5. Have the pupils write a letter to the principal on "Why I Want to Learn to Use Better English."

The test entitled miscellaneous B may then be used.

Do not tell the children what you are planning to do with the tests. Do not tell them what their score is on any of the tests. Do not tell them that they are going to have any more tests. Tell them that the practice exercises contain the correct forms. Do not answer a question about any of the items on the tests or practice sheets. If they ask you a question about it, tell them to wait and see.

6. Between test *miscellaneous B* and test verbs use the material called "Stowaways." Follow procedure outlined in No. 3.

7. Also between test *miscellaneous B* and test verbs use questions 3 and 4 on "Questions for Discussion" sheet, for an impromptu discussion taking fifteen to twenty minutes.

8. Between test verbs and pronouns use material called "Unconscious Humorists" following procedure outlined in No. 3.

9. Between test pronouns and test *miscellaneous A* use material on page 5 of the motivation material, following procedure outlined in No. 3.

The program of testing was identical with that used in the test motivation section of the experiment. The order of testing was reversed for P. S. 183, Brooklyn.

Samples of the material used are given below:

In an Electric Company

By J. W. Dietz, Western Electric Company

From the moment boys and girls join the Company, their command of English has a definite bearing on their advancement. In the remarks that follow, reference is made only to one group of New York offices.

Ordinarily, a boy's first job is the delivery of mail between the different buildings of the Company. His knowledge of English should be sufficient to enable him to secure any information he may need in his work, and to give a good account of himself when executives of the Company happen to talk with him. He must also be able to direct outsiders to buildings, departments and individuals. His English should be good enough to enable him to answer the questions of Company men and outsiders with accuracy, courtesy and without grammatical errors.

Clear speech is also essential here. When a boy gives a visitor or a Company executive a direct, courteous and clearly enunciated answer to a question he is creating a decidedly favorable impression for the Company and for himself.

This, of course, is equally true of girls. Young women enter the organization as filing clerks, as typists, as stenographers and as secretaries.

In their filing work, girls find fundamental English directly useful in the alphabetical and subject filing; and typists use their English training in correcting careless mistakes on written papers they have to copy, such as errors in spelling, or omission of words, et cetera.

Stenographers and dictating machine operators who have a good English background are able to transcribe letters without asking the dictators about paragraphing,

sentence structure or points of grammar. Secretaries are expected to have such command of English as will not only enable them to make certain that the correspondence leaving their desks is grammatical and accurate, but in addition enable them to compose their own letters and handle the less important mail themselves. When an executive hands a letter to a stenographer and says: "Please answer this and give them the information," he expects a well-written letter. Good English is important in the stenographer's job. It paves the way for her to increase her responsibility and is a distinct asset in becoming a secretary.

Young people coming with the Company should have the sort of background in English that will help them to improve their speech and writing as their work increases in responsibility. Where an employee has an interest in English that makes his outside reading a steady education, his chances of outdistancing his fellows are aided.

With girls and boys alike, good English is a valuable tool in mastering the details of a job and winning advancement. And the kind of English instruction that gives the student a taste for study of language and good literature gives him a real start toward success.

What has a definite bearing on the advancement of boys and girls in the Western Electric Company?

What kind of an answer should a boy be able to give to a visitor to the Company?

What jobs do young women take with the Company?

What is important in the stenographer's job?

What three things are secretaries expected to do?

QUESTIONS FOR DISCUSSION

Name situations where one would be at a disadvantage if he could not express himself clearly.

Suggestions: buying something in a store, ordering a meal in a restaurant.

Why is good English necessary in the work of a stenographer? lawyer? office boy? salesman? editor?

Do you always use correct English? Do your father and mother always use correct English at home? What are some of the things you can do to make your language more correct?

Have you heard any one mispronounce a word recently? Name a word you have heard mispronounced. Can you pronounce correctly the following: American, government, Tuesday, allies, hundred?

The other three passages and the questions for discussion which followed were similar in nature and content. The subjects, however, were different. One passage was entitled "Stowaways" in which the "stowaways" were the unconscious revelations of illiteracy often contained in letters. Another entitled "Unconscious Humorists," cited examples of ridiculous grammatical blunders. The fourth, presented in essay style, treated the inadequacy of the illiterate person in general situations.

One letter, typical of many written to the principal, is reproduced as evidence of the sincerity with which the pupils approached their task.

New York N. Y.
March 19, 1928.

Dear Mr. Theobald,

I would like to speak better english for it makes you a better American citizen. A good citizen alway's speak good english. This Tuesday I am going to get the address of the Good Speach Pioneers of America and I am going join that club.

Yours trully

By coincidence P. S. 165 was interested in a special campaign for better English in these very classes just prior to this experimental work. Certain children chosen by competition throughout the city were to be given the opportunity to speak before the radio. The elimination contest had already been held in this school so that interest in speech was motivated in a very real way as compared to the somewhat artificial methods employed in the experiment.

Results

The following table gives the gains on each test for each class.

Gains Made in Manhattan Schools

Room	Practice	Miscellaneous B	Verbs	Pronouns	Miscellaneous A
P. S. 87, No Motivation					
		One repetition	Three repetitions	Five repetitions	Ten repetitions
VIB1	−.94	−.56	.54	1.66	3.17
VIB2	1.42	.33	1.60	3.21	9.91
VIOpp	.60	−.58	.29	1.27	1.80
P. S. 10, Test Motivation—Three Repetitions of Practice Material for Each Test					
VIB1	.95	2.60	3.47	3.95	7.35
VIB3	−.16	3.21	3.04	5.88	6.83
VIB4	.11	1.33	2.66	3.75	4.23
P. S. 165, Intrinsic Motivation—Three Repetitions of Practice Material for Each Test					
VIB1	.92	.66	1.13	2.41	3.86
VIB2	.71	−.11	1.24	−1.05	.42
VIB3	−.18	1.11	.14	1.68	2.86

GAINS MADE BY BROOKLYN SCHOOLS

Room	Practice	Miscellaneous A	Pronouns	Verbs	Miscellaneous B
P. S. 114, No Motivation					
		One repetition	Three repetitions	Five repetitions	Ten repetitions
VIB1	.42	.34	.06	1.20	3.01
VIA3	1.55	2.52	1.75	.52	1.79
VIA4	.51	.36	.41	2.10	3.74
P. S. 210, Test Motivation—Three Repetitions of Practice Material for Each Test					
VIB1	.54	3.07	2.75	.83	.85
VIB2	−.28	4.01	.68	3.02	2.72
VIB3	−.54	3.06	.17	1.43	1.58
P. S. 165, Intrinsic Motivation—Three Repetitions of Practice Material for Each Test					
VIA1	2.63	−.50	1.46	.39	−.23
VIA2	−.67	−.20	−.15	−.27	−.36
VIA3	1.39	−.12	.34	1.46	−.02

These gains have been consolidated by weighted averages into the following table.

No Motivation

	One repetition	Three repetitions	Five repetitions	Ten repetitions
P. S. 87 Manhattan	− .31	+ .86	+2.09	+4.73
P. S. 114 Brooklyn	+1.07	+ .76	+1.33	+3.01
Test Motivation—Three Repetitions Throughout				
P. S. 10 Manhattan	+2.40	+3.08	+4.55	+6.30
P. S. 210 Brooklyn	+3.42	+1.21	+1.69	+1.57
Intrinsic Motivation—Three Repetitions Throughout				
P. S. 165 Manhattan	+ .62	+ .86	+1.31	+2.59
P. S. 183 Brooklyn	− .38	+ .39	+ .52	− .18

Finally a recombination of the groups gives the following weighted averages corrected by the table given on p. 22.

	Gains			σ_M	N
	Manhattan schools	Brooklyn schools	Weighted average		
No motivation, one repetition	− .31	+1.07	.496	.19	189
three repetitions	+ .86	+ .76	.794	.24	187
five repetitions	+2.09	+1.33	1.540	.26	179
ten repetitions	+4.73	+3.01	3.764	.35	185
Test motivation	+4.35	+2.08	3.028	.15	750
Intrinsic motivation	+1.66	+ .20	.692	.13	779

The gains on the first, second, third and fourth tests were computed for the four schools in the test-motivation and intrinsic motivation groups. Three repetitions of the practice material were included between each first and final test.

	Gain	N
First test	1.733	380
Second test	1.695	385
Third test	2.336	383
Fourth test	2.638	381

Conclusions

1. The most important single factor in learning as between repetition and the types of motivation used in this experiment is *amount of repetition*. Ten repetitions with only that minimum amount of motivation which habitually operates to cause pupils to do the accustomed work of the schoolroom and to carry out the instructions of the teacher caused more learning than any combination of three repetitions and more powerful types of motivation.

2. Intrinsic motivation such as manufactured by the devices used in this experiment caused no learning in addition to the practice carried on at the same time. There is strong evidence that motivation from an experienced situation is more effective than motivation from a described situation.

3. Test motivation caused learning over and above that which could be explained by practice. The value of the test motivation may be estimated as the equivalent of about five sheer repetitions.

4. There is evidence that there is more gain on tests given later in the series but it is not known whether this is due to the increased effectiveness of the motivation or to better adaptation to the practice.

Discussion

Many would deny that the motivation engendered by the readings, letters and discussions employed in the third section of the experiment is *intrinsic*. Intrinsic motivation means a more *intimate* and *seeking* interest than that probably aroused by the devices which we used. As usually described this intrinsic motivation originates during some activity when the way has been blocked and an obstacle arises which must be surmounted. It is doubtful whether any pupil was seeking employment during or previous to this experiment and had actually been turned down because of poor English; or had been unable to get food because he could not express himself; or had been made to feel uncomfortable by his friends because of his illiterate letters. The situations described and imagined by the pupils lacked the reality of experience. It is, however, significant to know that this kind of motivation cannot be conjured up by verbal description so as to have any effect. The actual intrinsic motivation which was found in P. S. 165 in the radio contest, although it probably came home to and was accepted as a challenge by only a few pupils was actually more effective than the purely described situations which were used in P. S. 183. This experiment probably will not shake the firm belief of many that intrinsic motivation is an important factor in learning. But our results do indicate that the mediate situations which come only by verbal description do not perform a motivating function vicariously. There is practically no transfer in motivation from a real to a verbally described situation.

That the test-motivation groups did show a gain over and above the gain which would come from the practice they had is significant. It indicates that competition with one's own past record, competition with other persons and competition as a member of a group with other groups are valuable types of motivation. A recent experiment by Sims[1] indicates that these three types operate in descending effectiveness. It is interesting to speculate how this motivation increases the learning. One obvious explanation is that it increased

[1] Sims, V. M.: The Relative Influence of Two Types of Motivation on Improvement. *Journal of Educational Psychology*, Vol. XIX, Oct., 1928, pp. 480–484.

the amount of practice. As the practice was conducted the sentences were read as a class orally and in unison. There was nothing to prevent pupils reading silently to themselves as the sheets were being distributed or collected or in the short interval between the class reading of one sentence and another. Indeed, motivation to be effective must make pupils reach out and seek practice and there is a possibility that some availed themselves of the opportunity to find out how to answer the test items correctly and hence to increase their score. That the motivation caused additional practice or "study" is an obvious explanation of the increased effectiveness of this type of motivation.

A second factor that may have operated is the analysis which the test motivation group may have made during the practice readings or during the testing. There is a possibility that some pupils may have noted differences between the errors in the first test in a series and the correct form in the practice material. Since certain of the constructions were repeated several times with varying context in test and practice material, similarities may have been noted. Again some pupils may have made an analysis of the grammatical principle involved while reading the practice sheets so that when the second tests were given they "applied the principle." Although this analysis is much less likely to be the explanation of the effectiveness of the test motivation some of the brighter pupils may have made effective use of analysis.

A third factor which may be used as an explanation is the quality of attention. If pupils read with greater attention to the form of the sentences and less attention to something else such as the content of the sentences this may have lended to the improvement on the tests. However it is not understood how this direction of the attention could influence the results except by means of some sort of an analysis as was described in the last paragraph.

Finally increased motivation may release certain inhibitions to learning and may make the three repetitions allowed more effective through a kind of facilitation. This, of course, is the usual explanation of how motivation operates. It is commonly assumed that it operates directly in making practice more effective. Further experimentations must be carried on to determine which of the factors mentioned is the real explanation of the additional gain due to the test motivation.

The question naturally arises as to the permanency of the learning resulting from the different types of motivation. On this point we

have no data. On the one hand it may be true that the same amount of learning, however attained, is equally permanent. On the other hand, many believe that learning with motivation is more permanent than non-motivated learning. *A priori*, there is good reason for believing that the non-motivated learning is more permanent. It is quite probable that many of the children *know* the correct forms but do not use them. Under motivation it is possible that they recalled and practiced what they knew. This learning would fall away to habitual performance when the immediate motive was removed. This matter needs experimentation.

Our findings with regard to the comparative ineffectiveness of motivation should be compared with the findings of other investigators, notably Knight and Remmers,[1] and Kitson[2] who find motivation decidedly potent. These studies, however, use learning functions which involve *speed* whereas our tests do not involve the speed factor. It is probable that motivation affects speed much more readily than the choice specific response itself.

Resumé

As to the practical implications of the experiment we must conclude that the most effective device that can be applied to learning is to increase the amount of drill or practice. The prime function of motivation is to make this drill or practice more palatable. No devices offer more hope for increasing learning than those which give each individual pupil more opportunity to practice. As to the most effective type of motivation this experiment points toward extrinsic motivation such as offering prizes, stimulating competition and informing the learner of his own progress as being most effective. Whatever the effectiveness of intrinsic motivation it is extremely difficult to organize in ordinary school procedure. Certainly in the matter of correct English one cannot have pupils attempt to obtain and hold real jobs, or deprive them of food or social companionship, or have their friends ridicule them for slovenly social letters until their English is grammatically acceptable. The school paper offers one real motivating situation but this reaches only a small proportion of pupils. This same difficulty applies to other subjects. The school

[1] Knight, F. B. and Remmers, H. H.: Fluctuations in Mental Production When Motivation is the Main Variable. *Journal of Applied Psychology*, Sept., 1923.

[2] Kitson, H. D.: A Study of the Output of Workers under a Particular Wage Stimulus. *University Journal of Business*, Nov., 1922.

cannot create a foreign-speaking community in which a foreign language can be acquired. Neither can it provide enough real situations involving algebra and geometry to provide for the learning of mathematics. One must be willing to postpone indefinitely or omit altogether much of the learning which we consider valuable in school if we have to depend on true intrinsic motivation. One cannot defend the position that unless a thing can be intrinsically motivated it is not worth learning at the present. Accepting this, children would not be taught to cross the street until they were knocked down by an automobile, or to brush their teeth until their teeth had decayed. Much of the efficiency and pleasure of living as social beings comes from learning things which can only be appreciated after they are learned.

The issue faced in the story of the boy who failed to learn "have gone" instead of "have went" is primarily one of transfer and not of learning. Perhaps the strongest argument in favor of intrinsic motivation is that the learning takes place in the situation in which it will be used. If greater learning takes place with isolated drill there is probably loss in transferring this learning to a real situation.

Finally described situations lack entirely the motivating force to facilitate learning. We must conclude that the situation that best provides for learning is one in which every pupil has abundant opportunity to practice under stimulations to which he is sensitive and in a situation containing as many elements as possible like those in which the learning will be used. Naturally some compromise must be made since it is practically impossible to obtain a maximum of these three conditions in ordinary school practice.

A NEW APPARATUS FOR PLOTTING AND A CHECKING METHOD FOR SOLVING LARGE NUMBERS OF INTERCORRELATIONS

L. DEWEY ANDERSON

Bureau of Educational Experiments, New York

AND

HERBERT A. TOOPS

Ohio State University

(*Continued from December*)

Part II. A Checking Method for Solving Intercorrelations

Upon the completion of the plotting of any one variable with other variables, the paper is removed from the board and the totals are entered in each small compartment in ink. This is unnecessary if N is not over one hundred fifty cases. The horizontal frequencies are entered in column V (see Fig. 1,[1] plot r_{12}). The sum of the X and Y marginal frequencies should be the same, that is, should equal the number of cases, or N. Since the vertical class intervals are indicated for all the plots at the same time, it is unnecessary to secure the vertical frequencies on more than one plot. Errors in the vertical plotting are exceedingly rare. The checking equations later used make the detection of any such misplotting a ready matter. *Approximately half the usual counting of marginal frequencies is thus saved.*

This method solves the correlations according to the following gross score formula:

$$r = \frac{(A + B - C)\frac{N}{2} - D \cdot E}{\sqrt{NA - D^2}\sqrt{NB - E^2}} \qquad (1)$$

Where, $A = \Sigma X^2 = \Sigma fd^2$ on x, where d is a positive step deviation from zero.

$B = \Sigma Y^2 = \Sigma fd^2$ on y.

$C = \Sigma(X - Y)^2$, which takes the place of the Σxy in the Pearson deviation method.

$N =$ Number of cases.

$D = \Sigma X = \Sigma fd$ on x.

$E = \Sigma Y = \Sigma fd$ on y.

[1] Figure 1 is to be found in the December instalment of the article.

This is the gross score adaptation of the usual Pearson r formula. In securing A, B, D, and E, a stencil is used. (See Fig. 2.) Deviations are measured from the 0 point in the distribution instead of the "guessed" average. This stencil is placed adjacent to the X frequencies and the d and d^2 of the stencil are multiplied by the corresponding frequencies. When a Monroe Calculating machine is used, the D, or E, totals (as the case may be) are summed on the two rows of keys at the left of the keyboard while the A, or B totals, correspondingly, are summed on the right hand side of the keyboard, both D and A, or E and B, being done in one operation. This process uses the calculating machine as two adding machines at one time. For example, if our step 1 frequency is 5, the 1 in the second column from the left on the keyboard and the 1 in the column at the extreme right are pressed down, and the crank is turned five times. If the frequency for step 2 is 4, 2 in the second column to the left and 4, that is, 2^2 (the d^2 for this deviation) in the extreme right column are pressed down, and the crank is turned four times, *i.e.*, as many times as there are frequencies. When all frequencies have been multiplied in this way, the left hand sum in the answer window gives D or E, and the right hand sum A or B. *As a check on this procedure the stencil is moved down one position so that the* 1 *deviation of the stencil takes the place formerly occupied by the* 0 *deviation,* while all other steps are similarly increased by 1. The same machine operations are then repeated and the following checks should hold:

If the new summations be represented by primed letters,

$$D' = D + N; \text{ that is, } \Sigma(X + 1) = \Sigma X + N \quad (2)$$

$$E' = E + N; \text{ that is, } \Sigma(Y + 1) = \Sigma Y + N \quad (3)$$

$$A' = A + 2D + N; \text{ that is, } \Sigma(X + 1)^2 = \Sigma X^2 + 2\Sigma X + N \quad (4)$$

$$B' = B + 2E + N; \text{ that is, } \Sigma(Y + 1)^2 = \Sigma Y^2 + 2\Sigma Y + N \quad (5)$$

Since A, B, D, and E are determined by the X and Y frequencies which are always the same, for any given variable considered, in all the correlation plots, *it is not necessary to resort to the above checks more than once,* if the marginal frequencies are always made to check.

The operations and checks up to this point, then, are to be done *only once for each variable.* Because of the checks involved in working C, it is not even necessary that the Y-frequencies be secured for more than the first time. It is necessary, of course, that the X and Y frequencies be identical on all the plots, but time is saved by correcting the plotting mistakes, if there are any, after these C checks are made.

d	d^2
18	324
17	289
16	256
15	225
14	196
13	169
12	144
11	121
10	100
9	81
8	64
7	49
6	36
5	25
4	16
3	9
2	4
1	1
0	0

FIG. 2.

d	d^2	
−17	289	
−16	256	
−15	225	
−14	196	
−13	169	
−12	144	
−11	121	
−10	100	
− 9	81	
− 8	64	
− 7	49	
− 6	36	
− 5	25	
− 4	16	
− 3	9	
− 2	4	
− 1	1	
0	0	
1	1	*

FIG. 3.

d	d^2	
0	0	
1	1	
2	4	
3	9	
4	16	
5	25	
6	36	
7	49	
8	64	
9	81	
10	100	
11	121	
12	144	
13	169	
14	196	
15	225	
16	256	
17	289	
18	324	*

FIG. 4.

FIG. 2.—Cardboard stencil for obtaining:

$$\Sigma X = D \text{ and } \Sigma X^2 = A; \text{ or, } \Sigma Y = E \text{ and } \Sigma Y^2 = B;$$

and

$$\Sigma(X + 1) = D' \text{ and } \Sigma(X + 1)^2 = A'; \text{ or,}$$
$$\Sigma(Y + 1) = E' \text{ and } \Sigma(Y + 1)^2 = B'.$$

FIG. 3.—Cardboard stencil for obtaining "above the diagonal" summations as follows:

"a. d." $\Sigma(X - Y)$ = "above diagonal" component of F;
"a. d." $\Sigma(X - Y)^2$ = "above diagonal" component of C;
"a. d." $\Sigma(X - Y + 1)^2$ = "above diagonal" component of T.

FIG. 4.—Cardboard stencil for obtaining "below the diagonal" summations as follows:

"b. d." $\Sigma(X - Y)$ = "below diagonal" component of F;
"b. d." $\Sigma(X - Y)^2$ = "below diagonal" component of C;
"b. d." $\Sigma(X - Y + 1)^2$ = "below diagonal" component of T.

---

* Multipliers of this box used only in checking.

This is particularly true when experienced workers are doing the correlations. If the two checks, soon to be described, hold in the case of every variable, it is not necessary to count either the X- or Y-marginal frequencies more than the one time when A, B, D, and E, are first obtained.

To secure $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$, it is necessary to obtain the diagonal frequencies. This is accurately done by means of the

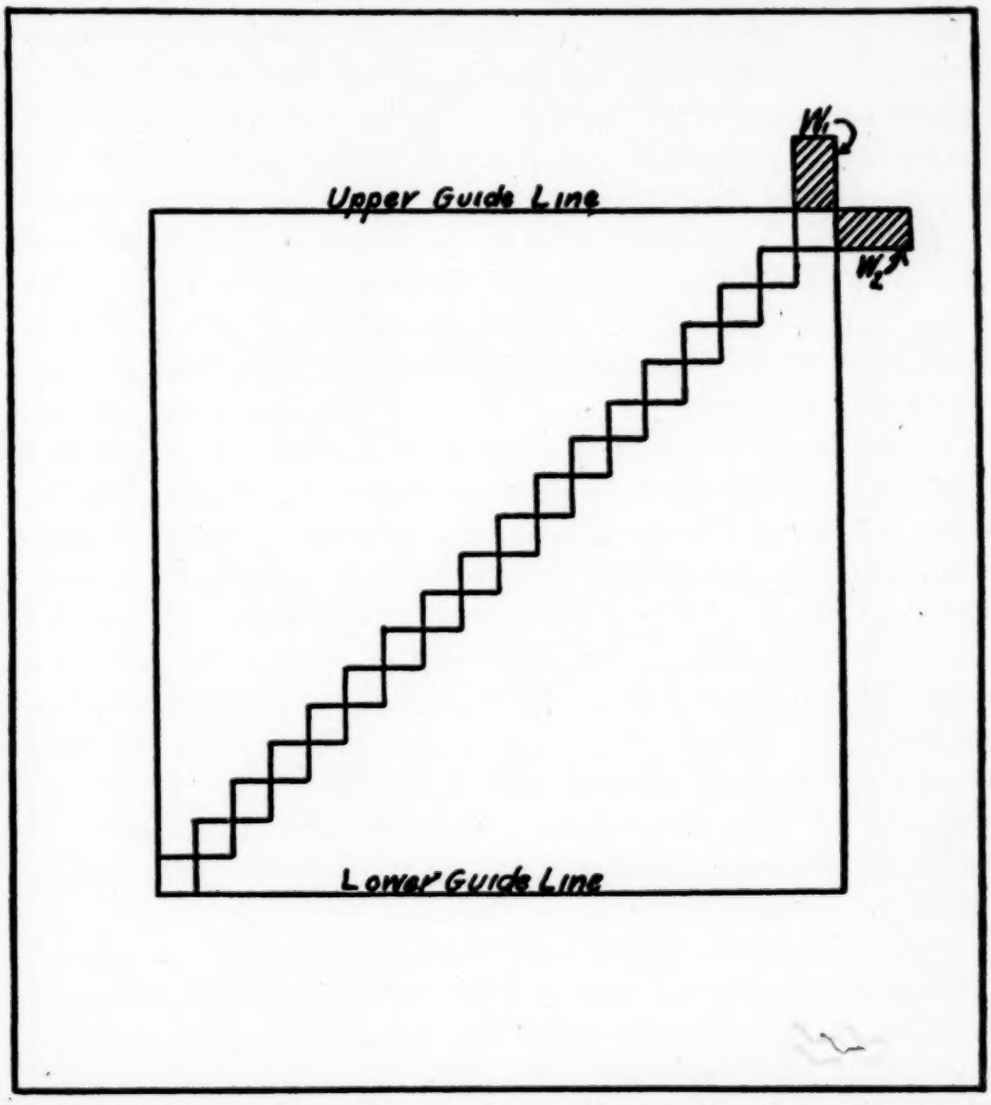


FIG. 5.—Celluloid diagonal frequency stencil used in counting diagonal frequencies.

celluloid stencil[1] for counting diagonal frequencies, Fig. 5. The stencil is placed on the data sheet so that the upper guide line is directly over the top horizontal line, and the lower guide line covers

[1] The lines are scratched into the celluloid; then inked in with black ink; and finally shellacked with clear shellac to make the stencil permanent.

the lowest horizontal line of the correlation plot. In other words, the stencil is placed so that the guide lines form the upper and lower boundaries of the correlation plot. The stencil is then moved horizontally until the tally marks in the diagonal row nearest the upper left hand corner of the data sheet appear in the small squares of the stencil. When the stencil is in this position the tally marks of a given diagonal row are counted and the total is entered on the data sheet through Window No. 1, W_1, on Fig. 5. The stencil is moved to the right one space at a time and the frequencies are similarly recorded until the tally marks in the longest diagonal have been counted and entered through W_1. From this point on the stencil is moved down (vertically) one space at a time and the tally marks of each successive diagonal row are added and are entered through W_2, (Fig. 5). The tally marks in the diagonal rows which are to the right of, and below, the longest diagonal row are thus counted in turn. These, (for r_{12}) are shown in column S and row S' of plot r_{12}, Fig. 1. The two squares indicated by *'s in this plot of Fig. 1 are always left blank.

To obtain both $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$, it is necessary to sum the extensions for those frequencies above the longest diagonal row (in the row S) and those frequencies below the longest diagonal row (in column S') independently. These are hereafter referred to as $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$ "above the longest diagonal;" and $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$ "below the longest diagonal," respectively. For obtaining $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$ "above the longest diagonal" the stencil, Fig. 3, is used. The 0 deviation is placed opposite the longest diagonal compartment and the d's and d^2's are multiplied by the frequencies. For obtaining the $\Sigma(X - Y)$ and $\Sigma(X - Y)^2$ "below the longest diagonal" the stencil shown in Fig. 4 is used.

The $\Sigma(X - Y)$ "above the longest diagonal" is obtained by ignoring the negative signs of Fig. 3 stencil, and then subtracting the sum, thus obtained, from the $\Sigma(X - Y)$ "below the longest diagonal." This algebraic total, $\Sigma(X - Y)$, as a check, should equal $\Sigma X - \Sigma Y$, or $D - E$.

If this does not prove to be true an error has been made and the operations should be checked in the following order: First, the extensions are recomputed to see that no errors were made; second, the X- and Y-frequencies are secured (this has not previously been done) and are then compared with the master X- and Y-frequencies; and

third, the diagonal frequencies are recounted. When the check holds, and only then, the two $\Sigma(X - Y)^2$ components are added to obtain C.

We now need to check C. The stencil for $\Sigma(X - Y)^2$, "above the longest diagonal" is moved up (to the left of the data sheet) one square to obtain the $\Sigma(X - Y + 1)^2$ component "above the longest diagonal." The stencil for the $\Sigma(X - Y)^2$ "below the longest diagonal" is moved up (towards the top of the sheet) one square to obtain the $\Sigma(X - Y + 1)^2$ component "below the longest diagonal." The two $\Sigma(X - Y + 1)^2$ are added; whereupon the following check should hold:

$$\Sigma(X - Y + 1)^2 = \Sigma(X - Y)^2 + 2\Sigma X - 2\Sigma Y + N,$$

or in letters,

$$T = C + 2D - 2E + N \tag{6}$$

To recapitulate, the following checks should hold, and should be applied in the order given:

Order	Check
(1)	$\Sigma(X + 1) = \Sigma X + N$
(2)	$\Sigma(Y + 1) = \Sigma Y + N$
(3)	$\Sigma(X + 1)^2 = \Sigma X^2 + 2\Sigma X + N$
(4)	$\Sigma(Y + 1)^2 = \Sigma Y^2 + 2\Sigma Y + N$
(5)[1]	$\Sigma(X - Y) = \Sigma X - \Sigma Y$
(6)[2]	$\Sigma(X - Y + 1)^2 = \Sigma(X - Y)^2 + 2\Sigma X - 2\Sigma Y + N$

If the fifth and sixth checks do not hold, the error resides either in the original plotting or in the counting of diagonal frequencies or in making the diagonal extensions and summations. The appropriate checks (1, 2, 3, or 4) should be applied to discover the source of the error or errors, and then the errors are corrected so that the checks hold, and any work subsequently done should be done over and checked.

---

[1] Master frequencies, *i.e.*, checked frequencies, become available after one has plotted all the twenty-four correlations with variable 1 and with variable 2.

[2] Checks 5 and 6 are, as above explained, the only ones regularly done on every one of the correlation plots if errors are not found.

As a final over-all check, we solve for r by the usual gross score formula and by a checking formula, (7) as follows:

$$r = \frac{\frac{N}{2}(A + B - C) - D \cdot E}{\sqrt{NA - D^2}\sqrt{NB - E^2}} = \frac{\frac{N}{2}(A' + B' - C) - D \cdot E'}{\sqrt{NA' - D'^2}\sqrt{NB' - E'^2}} \qquad (7)$$

$$r = \frac{\frac{N}{2}[\Sigma X^2 + \Sigma Y^2 - \Sigma(X - Y)^2] - \Sigma X \cdot \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{\frac{N}{2}[\Sigma(X + 1)^2 + \Sigma(Y + 1)^2 - \Sigma(X - Y)^2] - \Sigma(X + 1)\Sigma(Y + 1)}{\sqrt{N\Sigma(X + 1)^2 - [\Sigma(X + 1)]^2}\sqrt{N\Sigma(Y + 1)^2 - [\Sigma(Y + 1)]^2}}$$

These can best be done by systematic job analysis,[1] which checks independently even the final division which yields the correlation coefficient. Note particularly that the checking value, T, does not appear in the checking formula, but rather that C appears in the same location in both the original and the checking formula.

In using this method on a series of thirty-four variables, one hundred cases, the writers found the average time for solving each correlation to be twenty-two minutes. The following job analysis shows the time consumed for each of the operations:

TABLE II.—CORRELATION TIME ANALYSIS

OPERATION	MINUTES PER CORRELATION
Plotting time	4.9
Counting of diagonal frequencies	4.9
Checks on diagonal frequencies includes extensions and substitutions in $\Sigma\,\Sigma(X - Y) = \Sigma X - \Sigma Y$ and $\Sigma(X - Y + 1)^2 = \Sigma(X - Y)^2 + 2\Sigma X - 2\Sigma Y + [N]$	7.6
Adding and checking ΣX and ΣX^2 includes $\Sigma(X + 1) = \Sigma X + N$, $\Sigma(Y + 1) = \Sigma Y + N$, $\Sigma(X + 1)^2 = \Sigma X^2 + 2\Sigma X + N$, $\Sigma(Y + 1)^2 = \Sigma Y^2 + 2\Sigma Y + N$	.15
Working and checking numerators	3.15
Working and checking denominators	.15
Dividing numerators by denominators and checking	1.20
Total	22.05

The cost per correlation if the clerical assistants are paid forty cents an hour is *fifteen cents.*

[1] Toops, H. A.: Computing Intercorrelations of Tests on the Adding Machine. *Journal of Applied Psychology,* Vol. VI, No. 2, 1922, p. 177ff.

SUMMARY

A correlation plotting apparatus is described which permits rapid plotting of intercorrelations and increases the accuracy. The capacity of the apparatus is twenty-four correlation scattergrams from one set-up of the frame. A method of solving the correlations is described which not only reduces the number of calculations but also provides checking formulae at various steps in the process. A time analysis sheet on the plotting of thirty-four variables indicated that the total time per correlation ($N = 100$) was about twenty-two minutes.

A SIMPLIFIED METHOD OF COMPUTING MULTIPLE CORRELATION CONSTANTS

FRANK S. SALISBURY

Washington State Normal School, Bellingham, Wash.

At the outset of this article the writer feels the need of justifying its publication. This is due to the publication of an article in the September, 1926 quarterly of the American Statistical Association[1] which covers much the same ground as that intended herein, and which in addition has the advantage of certain important statistical refinements in procedure with all necessary statistical verifications. The present article seems however to be justified for the following reasons: In comparison the method here presented is less complicated and can be more readily mastered, and it constitutes a more convincing exposition of the relations which are operative in a multiple correlation problem, making it particularly desirable for instructional purposes. One who has used the simpler method will be the better able to appreciate the statistical refinements and verifications which Dr. Kelley has contributed to the method as it appears in the statistical quarterly. Both methods represent developments of his previous work to be found in his "Statistical Method."[2]

As to the range of usefulness of the method here presented it has been the experience of the writer that it serves quite as well as the refined procedure for problems employing as many as five or six variables, and after some acquaintance with the procedures but little difficulty has been encountered with most problems containing more variables.

The reader interested in the statistical verification of formulæ is referred to the article in the statistical quarterly. The exposition here will concern itself with the demonstration of a simple method for obtaining the Beta regression coefficients or weights to be given standard measures of the respective variables in the regression equation, and the multiple correlation coefficient.

Briefly stated, the essential steps of the method are as follows:

First, a weight is guessed for each of the independent variables.

---

[1] Kelley, T. R. and Salisbury, F. S.: An Iteration Method for Determining Multiple Correlation Constants. *Journal of the American Statistical Association*, September, 1926.

[2] Kelley, T. R.: "Statistical Method." The Macmillan Company, 1923, pp. 302–310.

Second, by a series of simple multiplications of weights and correlation coefficients between the independent variables and the summation of the products, data are obtained for judging the accuracy of the guessed weights.

Third, using these data in a simple key formula a series of corrections are made to the guessed weights, one at a time, always correcting the weight having the greater need, until no further corrections are necessary.

Fourth, the problem is checked for the accuracy of the computations made, and the multiple correlation coefficient is obtained after a few more computations of low order of difficulty.

Data for the demonstration of the method have been selected from a study of one hundred thirty-one freshman normal school students who, at the beginning of three months work in sight singing, were given a battery of tests of musical knowledge and abilities, among which were the five Seashore tests of musical ability—(1) Pitch, (2) tonal memory, (3) time intervals, (4) intensity, and (5) rhythm (the independent variables of the present problem), and a dictation test which was designed to measure the ability of students to listen to a series of measures and to write them upon staffs. The dictation test is the criterion of the present study, designated by the subscript "$_o$." The statistical problem may, for the present purpose, be stated as follows: We wish to find the most effective way of combining the five measures of musical ability into a composite score which may serve as a basis for predicting musical attainment as measured by the dictation test.

For the solution of this problem by the method here proposed we need the Pearson product moment coefficients of correlation between each independent variable and each of the other four independent variables, and the correlation between the criterion and each of the five independent variables. Having these basic data we proceed to determine the weights to be given the standard measures of the five independent variables so as to combine them into a composite measure which will have the maximum agreement with the dictation scores. The multiple correlation between the criterion and the composite is to be had as a measure of this agreement.

The basic data for the problem are to be found in the first six rows of Part I of Table I. In the first row are the r_{ou} correlations of the criterion dictation and the five Seashore tests—.538 with pitch, .454 with tonal memory, .565 with time, .201 with intensity and

.154 with rhythm. The entries of the next five rows are the r_{up} correlations between the five Seashore tests, the first row, for example, containing the correlations between pitch, variable #1, and the other four variables. The entry of the correlation "1.0" in each row is made to simplify computations to follow. As it appears in table it may be thought of in a reasonable way as the correlation of a given variable with itself.

Guessing the Initial Weights

The following rule, based upon experience, is suggested for determining initial weights with which to initiate the problem: For each variable use an initial weight that is approximately one-half the r_{ou} correlation with the criterion, of similar sign, and, to make computation easy and accurate, use the nearest one place decimal. The initial weights, the first five entries in the last column of the table, conform to this rule: .30 for pitch, .20 for tonal memory, .30 for time, .10 for intensity and .10 for rhythm. They are obtained by dividing the respective r_{ou} correlations of the first row by two and entering the nearest one place decimal. A more refined method of determining initial weights is given later.

Products of Guessed Weights and r_{up} Correlations

The next step is to compute the data for the five rows of Part II. The first row contains products obtained by multiplying the correlations of the r_{1p} row of Part I by the initial weight .30 of variable #1. Thus we multiply .30 by 1.0 to get .3000; by .485 to get .1455; .381 by .30 to get .1143; .320 × .30 gives .0960; and .202 × .30 to get .0606. In similar manner the correlations and weights of the last four rows of Part I are multiplied and their products entered in their respective places in the last four rows of Part II.

The next step is to add the five vertical columns of products that have just been computed, the sums being entered as the first row of Part III, labeled "First A-row." Let us examine the nature of these sums in terms of the addends. The first sum .5635 is obtained by adding .3000, .0970, .1143, .0320 and .0202. If this be written in equational form using algebraic symbols we would have,

$$A_1 \text{ equals } w_1 + (w_2r_{12} + w_3r_{13} + w_4r_{14} + w_5r_{15})$$

which may be written in a condensed form as follows,

$$A_1 \text{ equals } W_1 + \Sigma w_p r_{1p}$$

which may be generalized as follows,

$$A_u \text{ equals } w_u + \Sigma w_p r_{up}$$

TABLE I.—COMPUTATION OF REGRESSION COEFFICIENTS AND THE MULTIPLE CORRELATION

Symbols and titles	1	2	3	4	5	Initial weights
Part I						
r_{ou}	.538	.454	.565	.201	.154	
r_{1p}	1.0	.485	.381	.320	.202	.30
r_{2p}	.485	1.0	.394	.407	.511	.20
r_{3p}	.381	.394	1.0	.434	.420	.30
r_{4p}	.320	.407	.434	1.0	.251	.10
r_{5p}	.202	.511	.420	.251	1.0	.10
Part II						
w_1r_{1p}	.3000	.1455	.1143	.0960	.0606	
w_2r_{2p}	.0970	.2000	.0788	.0814	.1022	
w_3r_{3p}	.1143	.1182	.3000	.1302	.1260	Record of corrections
w_4r_{4p}	.0320	.0407	.0434	.1000	.0251	
w_5r_{5p}	.0202	.0511	.0420	.0251	.1000	
Part III						
First A-row	.5635	.5555	.5785	.4327	.4139	(#5) − .30
$-.30r_{5p}$	−.0606	−.1533	−.1260	−.0753	−.3000	
Second A-row	.5029	.4022	.4525	.3574	.1139	(#4) − .20
$-.20r_{4p}$	−.0640	−.0814	−.0868	−.2000	−.0502	
Third A-row	.4389	.3208	.3657	.1574	.0637	(#3) .20
$.20r_{3p}$	.0762	.0788	.2000	.0868	.0840	
Fourth A-row	.5151	.3996	.5657	.2442	.1477	(#2) .10
$.10r_{2p}$	.0485	.1000	.0394	.0407	.0511	
Fifth A-row	.5636	.4996	.6051	.2849	.1988	(#4) − .10
$-.10r_{4p}$	−.0320	−.0407	−.0437	−.1000	−.0251	
Sixth A-row	.5316	.4589	.5617	.1849	.1737	(#5 − .02)
$-.02r_{5p}$	−.0040	−.0102	−.0084	−.0050	−.0200	
Seventh A-row	.5276	.4487	.5533	.1799	.1537	(#4) .02
$.02r_{4p}$	.0064	.0081	.0087	.0200	.0050	
Final A-row	.5340	.4568	.5620	.1999	.1587	
Part IV (Check)						Final weights
$.30r_{1p}$	.3000	.1455	.1143	.0960	.0606	(#1) .30
$.30r_{2p}$	.1455	.3000	.1182	.1221	.1533	(#2) .30
$.50r_{3p}$	.1905	.1970	.5000	.2170	.2100	(#3) .50
$-.18r_{4p}$	−.0576	−.0733	−.0781	−.1800	−.0452	(#4) − .18
$-.22r_{5p}$	−.0444	−.1124	−.0924	−.0552	−.2200	(#5) − .22
Check A-row	.5340	.4568	.5620	.1999	.1587	
Final w_u	.30	.30	.50	−.18	−.22	
w_uA_u	.16020	.13704	.28100	−.03598	−.03491	σ_c^2 = .50735
w_ur_{ou}	.16140	.13620	.28250	−.03618	−.03388	s = .51004

The multiple correlation coefficient, $r_{oc} = \frac{s}{\sigma_c} = \frac{.51004}{.7123} = .7162$.

These A_u values play an important part in the procedure, being one of the members of the key equation. (The subscript "$_u$" indicates the independent variable under consideration at a given time, while "$_p$" indicates all the other independent variables taken in order.)

Correcting Weights by Use of the Key Formula

The next step is to make a series of corrections to the weights, one at a time, always correcting the weight having the greater need, until no further corrections are necessary. The key formula serves to identify the weight that should first be corrected, and indicates the amount of the correction to be made. The formula is, r_{ou} equals A_u, which equals $w_u + \Sigma w_p r_{up} \cdots$ key formula. By correcting the w_u weight of a variable the two members of the key equation are made to approximate each other, and this change likewise leads to a correction in the A-values of the other variables. These corrections to the A-values of other variables are not as great as the correction to the w_u weight, and the series of corrections proceeds to smaller and smaller corrections until the weights are of sufficient accuracy for their intended purpose.

Having the A_u values, first A-row of Part II, and the r_{ou} correlations in the first row of Part I, it is easy to make the comparisons of these values for each variable and discover the variable which shows the greater difference. Following this method, which variable needs the greater correction? The greatest difference is found in #5, the A_u value .4139 requiring a reduction of approximately .26 to make it equal to r_{ou} .154. However, instead of making the correction of a $-.26$ it is much easier and quite as effective at this stage of the problem to use a correction which is the nearest one-place decimal. This rule may well be followed until the corrections have become as small as .05 or less. To make the correction of $-.30$ to the weight of #5, and to the $w_p r_{up}$ values of the other variables, it is necessary to multiple the r_{5p} coefficients of the sixth row of Part I by the correction $-.30$, enter the products in their respective columns below the first A-row values, and by algebraic addition obtain the second A-row values. These new A_u values are the same as would be obtained had the original weight of #5 been $-.20$ instead of .10.

With the second, A-row computed, the new A_u values are in turn compared with the respective r_{ou} coefficients in the first row of Part I to discover the variable which now shows the greater need of change. The greater correction is demanded by #4, $-.16$ being the approximate correction needed to make the A_u value .3574 approximate closely the r_{ou} .201. Therefore we correct #4 by $-.20$, the nearest one place decimal, by multiplying the r_{4p} coefficients of Part I by the correction $-.20$, entering the products below the second A-row values to which they are added algebraically to obtain the A_u values of the third A-row. Attention is called at this point to the method of

identifying the products by the symbols of the first column and to the recording of corrections in the last column.

Repeating the process, it is found that #3 is to have a correction of a plus .20, the nearest one place decimal to the difference between A_u and r_{ou}. The $.20r_{3p}$ products are obtained, entered, and the new A_u values of the fourth A-row computed. The next correction is to #2, the nearest one place decimal being plus .10. The $.10r_{2p}$ products are entered and added to the values of the fourth A-row to make the fifth. Variable #4, through additions to its A_u value from the corrections of #3 and #2, calls for a correction of approximately − .10. This correction being made, the resulting sixth A-row values show a fairly close agreement in the two members of the key formula, the greater need for correction being found with #5, the amount being only −.02. This correction made, the seventh A-row values show #4 demanding a correction of plus .02, at which point the A_u values of the final A-row are found to agree with the respective r_{ou} values to the nearest two place decimal. Further corrections are possible, but for practical purposes the corrections may be considered as completed since further refinement of weights will have negligible effect upon composite scores.

Summing Corrections—Final Weights

It is now necessary to add to the initial guessed weights the various corrections that have been made. The following table shows this summary, the data being taken from the last column of Table I in which the initial weights and the corrections for the respective variables have been recorded as the problem has progressed.

Table II.—Computation of Final Weights

Variable	Initial weight	Correlations	Final weight
#1	.30	(None)	.30
#2	.20	+.10	.30
#3	.30	+.20	.50
#4	.10	−.20 −.10 +.02	−.28
#5	.10	−.30 −.02	−.22

Checking for Accuracy

Part IV of Table I shows the data computed for checking the accuracy of the computations that have been made. The final weights of the five variables are multiplied into the respective r_{up} coefficients of Part I to secure the data for computing the "Check A-row," a

procedure paralleling that employed in obtaining the data of Part II. The resulting A_u values are found to be essentially in agreement with the values of the final A-row of Part III, indicating that all computations have been correct. Should this check reveal errors the A_u values of this check may, of course, be used as a basis for necessary corrections in continuing the problem.

The Multiple Correlation Coefficient Computed

To compute the multiple correlation coefficient the A_u values of the "Check A-row" are multiplied by the respective final weights to obtain the $w_u A_u$ products, .16020, .13704, etc. These products are added to obtain the σ_c^2, equal to .50735, which appears in the last column. Then the r_{ou} coefficients of the first row of Part I are also multiplied by the respective final weights to obtain the last row of the table containing the $w_u r_{ou}$ values, which are summed to give "s," equal to .51004, which is entered in the last column. The formula for the multiple correlation is

$$r_{oc} \text{ equals } \frac{s}{\sigma_c} \text{ equals } \frac{\Sigma w_u r_{ou}}{\sqrt{\Sigma w_u A_u}}$$

Using this formula the multiple correlation has been computed at the foot of Table I and found to equal .716.

Whence the Key Formula

The origin of the key formula is due to the observation that, as the series of weight corrections bring the problem nearer and nearer to completion, the s and σ_c^2 values come to approximate each other very closely. These values being summations, inquiry led to the discovery that the respective addends were likewise tending to approximate each other. Later it was observed that by factoring out the w_u weight from these two values we would have the simple guide to the correction of weights in the fact that the A_u value should come to equal the r_{ou} coefficient. The record of this development may be written in the following equations:

$$s = \sigma_c^2$$

which becomes by substitution of equivalent values

$$\Sigma w_u r_{ou} = \Sigma w_u A_u = \Sigma w_u \ (w_u + \Sigma w_p r_{up})$$

and for any one variable, the equation for the respective addends would be,

$$w_u r_{ou} = w_u A_u = w_u \ (w_u + \Sigma w_p r_{up})$$

from which w_u may be factored to give

$$r_{ou} = A_u = w_u + \Sigma w_u r_{ou} \text{ (key formula)}$$

Obtaining Better "Initial Weights"

For the purpose of estimating initial weights it may be of advantage to rewrite the key formula as follows

$$w_u = r_{ou} - \Sigma w_p r_{up}$$

from which it appears that the choosing of a reasonable initial weight is a matter of forecasting a reasonable $\Sigma w_p r_{up}$ value, which may be subtracted from the r_{ou} correlation with the criterion to obtain the weight. The $\Sigma w_p r_{up}$ value, however, calls for consideration of the r_{up} correlations with the other variables and presupposes reasonable weights for each of those variables. This last fact brings us back where we started with the only profit that of pointing out the determining elements of the problem of choosing initial weights, namely, (1) the r_{ou} correlations with the criterion; (2) the correlations with the other independent variables; and (3) the weights which these variables are apt to have. With these factors in mind, and with some experience to acquaint one with ways of making this analysis useful, it will be possible to guess weights which will be better than those obtained by taking approximately one-half the correlations with the criterion.

In addition to this assurance of increase in skill in guessing initial weights the following definite procedure is recommended for making a series of rough corrections to guessed weights before they are adopted for use in Table I. This procedure is based upon the use of the key formula, the A_u values being approximated through the summing of roughly computed $w_p r_{up}$ products. To get these products a new order of making the computations is introduced. Turning to Table I, we will follow the routine of this new attack.

1. Instead of entering the initial weights in Part I they are to be placed on the edge of a separate piece of paper in a column order so that they can be placed adjacent to the r_{up} correlations of any desired column of Part I.

2. Beginning at the right of the table, roughly estimate the product of the respective weights and the r_{up} correlations, enter the products upon a piece of scrap paper and add them together to obtain an approximate A_u value. The computations are indicated below for variable 1.

Weights ×	r_{up}	Equal about
.30	1.0	.30
.20	.485	.10
.30	.381	.12
.10	.320	.03
.10	.202	.02
Approximate A_u is		.57

3. Comparing this approximate A_u of .57 with the r_{ou} .538 the necessary correction to the weight of #1, −.03, is obtained and recorded near the guessed weight of #1.

4. Repeating the procedure, the column of weights are moved to the next column of r_{up} correlations for #2, the multiplications are made to obtain the approximate $w_u r_{up}$ products (.15, .20, .12, .04, .05), which summed give an approximate A_u of .56. This is about .11 greater than the r_{ou} .454, and a correction of −.11 is to be entered near the guessed weight for #2.

5. The procedure is repeated for the other three variables, at which time it will be possible to determine an order in which corrections may best be made to the guessed weights. The weight needing the greatest correction is erased and a new weight (nearest one place decimals being used) is put in its place. The new weight is estimated by an algebraic addition of the correction to the guessed weight. It is possible that more than one weight should be changed at this point, depending upon the size of the corrections.

6. Having a new set of weights the procedure is repeated as outlined above, except that from this point the weights are changed as soon as the correction is determined. The procedure may well continue until the corrections have been reduced in size to less than .10.

7. Having the weights within reasonably accurate bounds, they should be entered in the last column of Part I and the method demonstrated in Table I followed till the problem is completed.

Other Constants

Other constants of the multiple correlation problem need not be treated here. Having the weights, or Beta regression coefficients and the multiple correlation coefficient, the other constants can be gotten by using formula to be had in any text dealing adequately with the multiple correlation problem. This statement will hold for the regression coefficients to be used in regression equations using gross measures, the alienation coefficient, and standard and probable errors of estimate. For the computation of partial correlation coefficients the reader is referred to the article in the Statistical Quarterly to which reference has already been made.

CONCERNING THE STANDARD ERROR OF A DIFFERENCE

HELEN M. WALKER

Teachers College, Columbia University

The importance to educational research of the technique of measuring the standard error of a difference is suggested by the fact that in 1926 and 1927 there appeared in the *Journal of Educational Research* and the *Journal of Educational Psychology* alone eleven papers in which the argument was based either entirely or in large part upon the size of the standard error[1] of a difference. These papers relate to such subjects as the superiority of one group, one race, one sex, one method of teaching, or one method of scoring tests, over another, and to the gain or loss made by pupils under different experimental procedures. The authors of these papers have been careful to state the reliability of the observed differences in terms of the standard error or the probable error of those differences. In addition to these eleven papers, there are in the same journals a very large number of other studies in which the chief point at issue is the reality and the importance of a difference, but in which no reference at all is made to the reliability of that difference or to the probability that it might have arisen through the operations of chance alone.

Since so many studies are basing conclusions of considerable practical import upon the reliability of a difference, it seems profitable to discuss the correct method of measuring that reliability, the distinction between the two formulas for the standard error of a difference, and the circumstances which justify the use of the shorter formula. The writer ventures the opinion that in at least half of the educational studies utilizing this technique, the shorter formula has been used where there is no warrant for it, with the result that the reliability of the difference, as reported, is incorrect. The differences usually studied in the field of education are of such a nature that the use of the short formula makes the standard errors appear too large, and therefore makes the differences appear less reliable than they actually are. If an investigator, using the short formula, has decided that a certain difference is a significant one which cannot reasonably be ascribed to chance, his conclusions would in all probability be strength-

[1] Whatever is said in this paper concerning the use of the standard error of a difference will obviously apply with equal force to the use of the probable error.

ened by the use of the correct formula. If an investigator, arguing from the abbreviated formula, has decided that a particular difference is so small in relation to its standard error that it cannot be considered significant, it is quite possible that the use of the longer formula would cause him to reverse his conclusions. This would not be true if there were a negative correlation between the variates contributing to the difference, but such cases are rare in educational research. Thus in almost all the studies in which the educator is likely to be interested, the use of the longer formula will have the effect of making the difference appear more convincing.

In the eleven papers mentioned, the standard errors or the probable errors of two hundred eighty-seven differences are stated. Only one of the authors specifically states the formula by which the reliability of the difference is computed. Six others give enough to make it possible to ascertain the method of computation. The remaining four furnish no clue to indicate which formula has been used. In each of the seven papers where it is possible to discover the method of computation, it is the shorter formula which has been used. In some of these studies the correlational term occurring in the longer formula could not have been found and therefore the longer formula was out of the question; in others it seems reasonable to assume that $r = 0$; but in at least six out of the eleven studies the correlational term could have been found and ought to have been used. What is still more significant, not one of the writers of these papers makes even passing reference to the possible existence of correlation between the variates contributing to these differences.

All who are interested in the theory of statistics know that the correct formula for the standard error of a difference is

$$\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2 - 2r_{xy}\sigma_x\sigma_y}$$

and that this can be shortened to

$$\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

if and only if $r_{xy} = 0$. Unless it is reasonable to assume that x and y are independent traits, or at least uncorrelated traits, then the use of the shorter formula cannot be justified, and will give rise to erroneous conclusions.

In many of the studies employing this technique, the difference in question is a difference between two means, and the formula becomes

$$\sigma_{M_x - M_y} = \sqrt{\sigma^2_{M_x} + \sigma^2_{M_y} - 2r_{M_xM_y}\sigma_{M_x}\sigma_{M_y}}$$

which is equal to

$$\sigma_{M_x - M_y} = \sqrt{\frac{\sigma_x^2}{N_x} + \frac{\sigma_y^2}{N_y} - 2r_{xy}\frac{\sigma_x\sigma_y}{\sqrt{N_x N_y}}}$$

if the assumptions of random sampling have been met. It is known[1] that the correlation between two sets of means is the same as the correlation between the original scores when the means are the means of random samples. This often makes it possible to utilize the correlation between original scores, a correlation which can be obtained, instead of the otherwise unknown correlation between two series of inaccessible means. This will be illustrated in more detail in the subsequent portion of this paper.

Thus the theoretical argument is extremely simple and unequivocal in its conclusion: The short formula can be justified only when $r = 0$, and any writer who employs the short formula is under the logical necessity of showing the reasonableness of his assumption that the two variates are uncorrelated. Practical considerations, on the other hand, may present serious difficulties. All too often the correlation which should be used is unobtainable. It is, however, available somewhat more often than the practices of investigators would imply. It has been said that the differences in question are sometimes differences between scores and sometimes differences between means. We will now discuss certain typical situations where the differences—whether between scores or means—are:

(*a*) Differences between the performances of the same individuals on two occasions,

(*b*) Differences between the performances of equated or of related individuals,

(*c*) Differences between the performances of equated or of related groups, and

(*d*) Differences between the performances of unrelated groups.

(*a*) *Differences between the Performances of the Same Individuals on Two Occasions.*[2]—This is the simplest of all the cases and provides an admirably clear illustration of the general principle. Let x represent a score on a final test and y a score made by the same person on an

[1] See Kelley: "Statistical Method" p. 178, formula 118.

[2] The discussion by Kelley on Differences in Intelligence and Achievement Scores. in the *Journal of Educational Psychology*, Vol. XIV, 1923, pp. 3213–34, and Holzinger's Note on Professor Kelley's Formula for Determining the Significance of Differences in the same journal, Vol. XVI, 1925, pp. 48–51, relate to one special aspect of this general case.

initial test. Then $x - y$ is the gain or the loss, the measure of the change effected by whatever experimental procedure has taken place between the time of the two tests. In this case r is quite obviously the coefficient of correlation between the two sets of scores made by the same persons, and can be found if enough individuals are measured to make it worth while to compute a correlation. Obviously this correlation is almost never zero and seldom negligible. It should never be overlooked in this case. If some reader is of the opinion that one formula will serve as well as the other, we recommend to him a very simple experiment. Let him obtain the scores made by the same persons on two different tests, or if real data are not available, let him construct two series of hypothetical scores such that there is a definite correlation between the two series. Then let him construct a third series—the gains—by subtracting each score on the first test from the corresponding score on the second test. Now let him compute the standard deviation of each of these three series. If the square of the standard deviation of the series of gains is equal to the sum of the squares of the standard deviations of the series of original scores, then the use of the short formula is vindicated for this case. If, however, there proves to be a discrepancy, let the experimenter compute the coefficient of correlation between the first and second series of test scores, and let him substitute this for r in the longer formula. If now the standard deviation as computed directly from the series of gains agrees with the value given by the longer formula, *i.e.*, $\sqrt{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}$, then the latter has been vindicated for this case. If neither formula checks with the direct computation, the computer may profitably subject his arithmetic to scrutiny!

Whenever the difference under consideration is the difference between initial and final scores on a test, or the difference between the work of the same persons under varying circumstances, then there is no justification for using the shorter formula and to do so produces an erroneous result. This is equally true when the question relates to the difference between the mean score of a group on an initial test and the mean score of the same group on a final test, and in this case the correlation between the scores of individuals on the two tests will furnish the value of r. If the two tests could be given again and again to many groups of children, and if a correlation could be computed between the two mean scores made by each group, this correlation between means might be expected to agree with the correlation between scores within the limits of sampling.

(*b*) *Differences between the Performances of Equated Individuals.*— Suppose that an investigator has selected one hundred pairs of children matched according to CA, MA, scores on a reading test, and on as many other counts as he pleases. Then for a stated period he subjects the two equated groups to widely different types of training, and at the end of the period gives both groups another form of the reading test. Now let us suppose that he finds the mean score of group A to be 3.5 points higher than the mean score of group B, and finds the standard deviations of the two groups to be 15 and 16. The investigator argues that the standard error of the difference of the two means is $\sqrt{\frac{(15)^2}{100} + \frac{(16^2)}{100}} = 2.19$ and that therefore the probability that a difference as large as 3.5 or larger, and in the same direction, should arise by chance where no real difference exists, is .055, and he decides that the magnitude of the difference is not very convincing. He is wrong. He should correlate the score made by each child in the first group with the score made by the other member of the same pair in the second group, and should use this for r in the longer formula. For purposes of illustration, let us suppose that the coefficient of correlation thus found is .60. Then the standard error of the difference between the means is not 2.19 but $\sqrt{\frac{(15)^2}{100} + \frac{(16)^2}{100} - \frac{2(.60)(15)(16)}{100}} = 1.39$. The probability that a difference as large as 3.5 or larger, and in the same direction, should arise by chance where no real difference exists, now appears to be only .006. Odds of six chances out of one thousand are considerably more convincing than odds of fifty-five chances out of one thousand, and have the additional advantage of being correct.

If the children have been well equated, the correlation may be fairly high, as the two members of the same pair are likely to occupy similar positions in the two series of scores. Certainly the correlation will seldom be negligible. This correlation will have as a practical upper limit the coefficient of reliability of the test employed, for it is hardly to be expected that the relationship between the scores of equated individuals will be higher than the relationship between the scores made by the same individuals on two forms of the test. *Whenever individuals have been equated in sufficient numbers to justify the computation of a coefficient of correlation, the longer formula should be used.* Of course, if only a few individuals have been equated, no

reliable value of r can be found, but under such circumstances all the measures computed must be considered unreliable.

A somewhat similar case arises when it is desired to compare the measures of related individuals, as, for example, some trait of a group of boys with the same trait of their sisters, or the scholastic achievements of two students selected as being at the fortieth and the sixtieth percentiles of their groups on a mental test. Here the r is probably not so great as the r found between equated individuals, but it is not zero by any means.

(*c*) *Differences between the Performances of Equated Groups.*—Of all four cases, this is the only one which presents serious difficulties. Suppose that two groups have been so chosen that their mean CA, mean MA, and mean score on an initial test, and their standard deviations on these various counts, are alike. After some experimental procedure, a final test is given to each group. Now we do not have equated individuals, and therefore cannot find r as in the case previously discussed.

It is necessary to consider the conditions of sampling under which other pairs of equated groups might be chosen. If it is understood from the conditions of the problem that all later pairs are to be so selected that their initial scores on the given test will be equal to that of the first pairs, then there is no reason to assume any correlation between the pairs of mean scores on a final test, and the short formula is entirely applicable. If, however, it is understood by the experimentor that later pairs are to be equated merely to each other and not to the first sample, then some correlation would very probably exist between the pairs of final mean scores. The problem is essentially that of the manner of definition of the population from which samples are to be drawn. If this population is defined as "all possible groups having an initial mean score equal to that of the groups already measured" then it is reasonable to assume independence between mean scores on a second test. If, however, the population is defined as "all possible pairs of groups equated for initial scores on the given test," the results might be somewhat like the following:

	Mean score on first test		Mean score on second test	
	A	B	A′	B′
First pair of groups	75.2	75.2	81.5	83.2
Second pair of groups	66.6	66.6	74.5	76.7
Third pair of groups	71.0	71.0	78.1	77.9
Fourth pair of groups	69.4	69.4	76.6	79.3
and so on.				

Under these circumstances there is likely to be a correlation between A′ and B′ which should be recognized in a discussion of the standard error of their difference. If one of the two equated groups makes a particularly high mean score, the other group is likely to have a fairly high record. If one of a pair of groups is of poor ability, the other has little chance of making a high mean score. Yet the actual value of the correlation is utterly inaccessible because we have only one pair of groups, not a series of pairs. In this case there seems to be no way of discovering what the standard error of the difference of the means really is. The fact that r is inaccessible does not make it zero and does not make the use of the short formula yield a trustworthy result. The use of the short formula provides an upper limit for the standard error of the difference when the correlation between the means is positive, or a lower limit if it is negative. If the coefficient of reliability for the test is substituted for the evasive r, then the long formula provides a lower limit for the standard error if r is positive, and an upper limit if r is negative. Thus, at least, limits may be set for the true value, which is probably somewhere between them and probably not coincident with either one. It seems to the writer very much better to admit that the true standard error of the difference has not been found, but that it lies between certain limits, than to assert that the upper limit is the true value. The latter assertion is based on the untenable assumption that because data are lacking for the computation of r, then r is necessarily zero.

Sometimes a research worker is concerned with a difference between the means of two groups which have not been equated but which are related in some definite fashion, as if they are the upper and lower fourths of the same group, or the upper and lower tenths. If the group as a whole is superior the scores made by both sections are likely to be higher than the scores made by the extreme fourths of a very poor

group. There is likely to be some correlation between the mean scores of the two sections, but probably not as much correlation as between the mean scores of equated groups.

(*d*) *Differences between the Scores Made by Two Unrelated Groups.*—This is the only case in which the writer would use the short formula without apology. For example, suppose a random sample of Grade VIII children in rural schools is to be compared with a randon sample of Grade VIII children in city schools, it seems entirely reasonable to make the assumption that, if sample after sample should be taken, there would be no correlation between the mean scores of pairs of samples.

Recommendations

1. Whenever it is possible to determine the value of r, the longer formula should be used. In particular, whenever the differences are between two performances of the same individual or between the performances of two related individuals, the correlation should be computed and used.

2. Whenever an investigator thinks the use of the short formula is justified on the ground that the groups are independent, he should show the reasonableness of the assumption that $r = 0$.

3. Whenever a difference must be found between two related groups where the value of the correlation is truly inaccessible, the writer should call attention to the fact that the standard error as found by the shorter formula is not a real value of the standard error but merely an upper limit. If possible he should also indicate a lower limit, as suggested under (*c*).

4. In any study utilizing the standard error of a difference, the writer would do well to indicate which formula he has used, in order that the reader may know how to interpret the author's conclusions.

GRAPHICAL COMPARISON OF TWO FORMULAS FOR THE STANDARD ERROR OF A DIFFERENCE

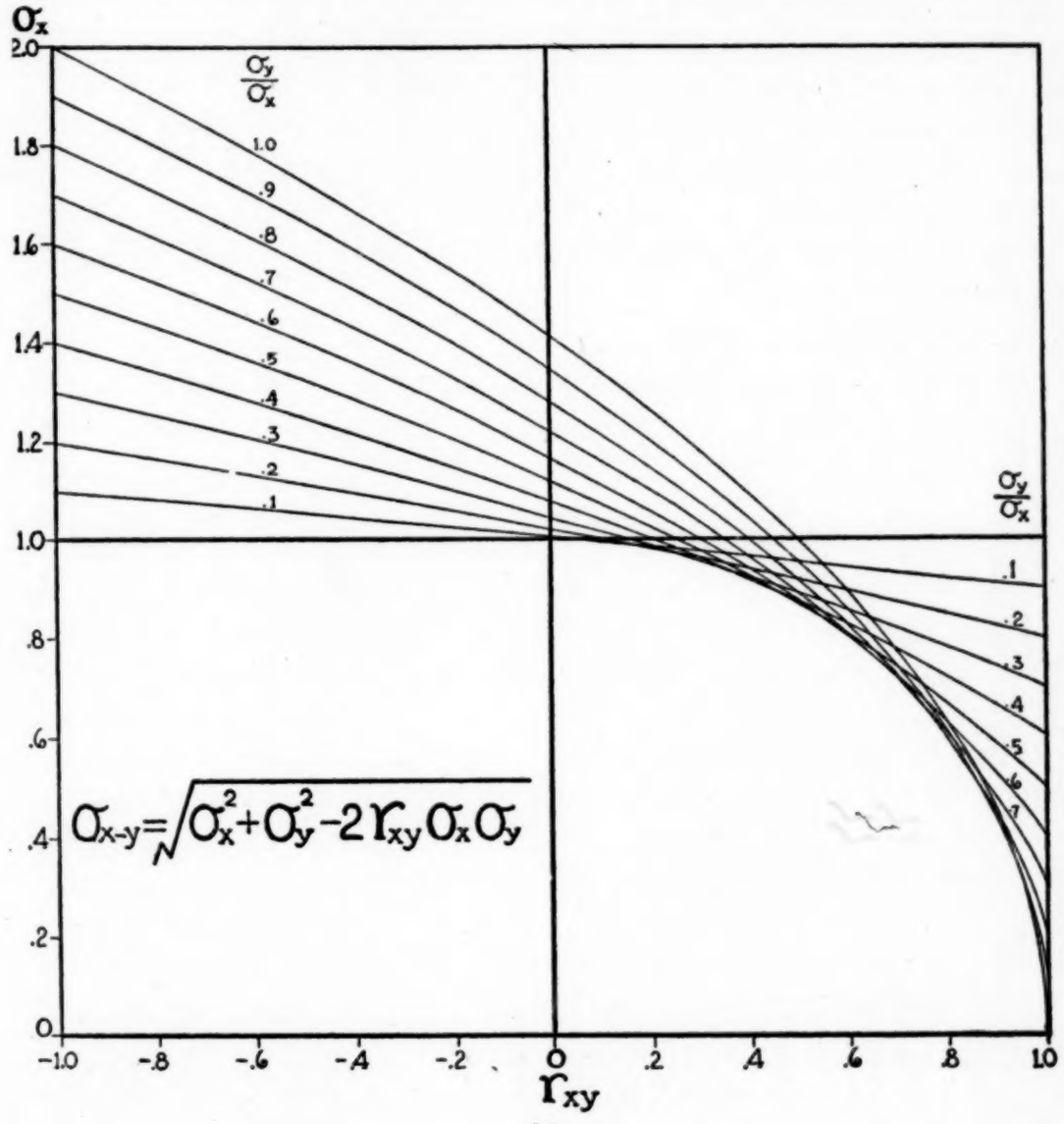
DOROTHY C. ROWELL

Teachers College, New York

The accompanying graph shows values of σ_{x-y} for ten values of the ratio $\frac{\sigma_y}{\sigma_x}$, at all values of r_{xy} from -1 to $+1$. Comparison is made between values found by the complete formula,

$$\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2 - 2r_{xy}\sigma_x\sigma_y},$$

and those found by the "short formula" often used, $\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2}$, which is the special case of the longer formula in which



$r_{xy} = 0$. Since $\sigma_{x-y} = \sigma_{y-x}$, one may arbitrarily take as a unit whichever happens to be larger, σ_x or σ_y. Measuring σ_{x-y} in terms of σ_x, the formula becomes $\frac{\sigma_{x-y}}{\sigma_x} = \sqrt{1 + \frac{\sigma_y^2}{\sigma_x^2} - 2r_{xy}\frac{\sigma_y}{\sigma_x}}$. Note that values for the "short formula" are represented by intersections of the central vertical line at $r_{xy} = 0$ with the complete curves, and that they are respectively less than, or greater than those of the longer formula, according as r_{xy} is negative or positive.

It is clear that as r_{xy} increases in absolute value, and as σ_y approaches equality with σ_x, the difference between the formulas becomes more and more significant. It is only in the rather unusual case in which σ_y is a very small fraction of σ_x that the correlational term may be negligible. In the common case of measuring gain by using two forms of the same test, r_{xy} is presumably large and positive, while σ_y will approximately equal σ_x. For example, if $r_{xy} = .5$, then σ_{x-y} cannot exceed σ_x. If $r_{xy} = .75$, unless σ_y is less than $.5\sigma_x$, σ_{x-y} will lie near $.7\sigma_x$. If $r_{xy} = .9$, and $\sigma_y = \sigma_x$, σ_{x-y} is less than $.45\sigma_x$. Under such circumstances, σ_{x-y} will be significantly smaller, and the reliability of the observed difference greater than the short formula would indicate.

THE DIFFICULTY OF AN ITEM

ELISHA F. BLISS, JR.

Plainfield, N. J.

In this article the writer will define the difficulty of an item or single task, derive a formula for it, and then compare it with other definitions of difficulty which have been used or proposed.

It may be considered that the scale upon which ability is measured is infinite, and that the difficulty of an item is represented by a point on this scale. Individuals whose ability is greater than that represented by this point will tend to pass this item while individuals whose ability is less will tend to fail.

The validity of an item is the extent to which the item differentiates between individuals of superior and inferior ability.

It will be seen that while we have defined the broader concept of validity, in practice we must define an item's validity in terms of the variability of the group upon which it is computed.[1] The difficulty of an item, however, is neither a function of the midpoint nor of the variability of the group upon which it is computed, but depends upon the units of the scale.

When the analyses of several items are based upon the same group, it is convenient to define the difficulty of an item as the distance above or below the mean of that group of a point which would be the midpoint of a normal distribution fifty per cent of whose population would pass the item.

From the definition of difficulty, it is obvious that those cases whose criterion of ability is that of the difficulty of the item will be equally divided between passes and failures. As we go up the scale the proportion of passes to failures increases, and as we go down the scale this proportion decreases.

Let us now consider the double frequency distribution, the correlation between criterion of ability and score on an item, which represents the validity of the item.

Let X be the mean of the criterion scores, and Y the mean of the item scores.

Let Q be the line cutting off q proportion of the cases and differentiating the passes from the failures. Q is perpendicular to X (given).

[1] In this, and the following discussions, it will be understood that the statements are true, subject to the probable errors of the measures involved.

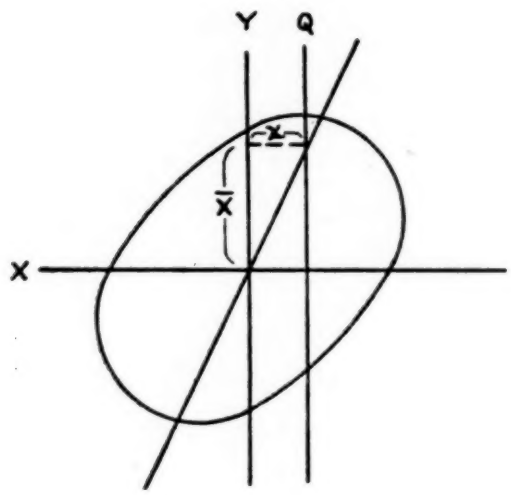
Then Q will intersect the regression line passing through the means of the arrays of item scores at a point which is $\overline{X}$ removed from the mean of the criterion.

Since the proportion of passes to failures in the array of which this point is the mean is equal to one, or corresponds to that which is found at the mean of the theoretical distribution, the distance $\overline{X}$ satisfies the definition of difficulty.

The regression line which fits this definition of difficulty is,

$$x = r\frac{\sigma_2}{\sigma_1}\overline{X}$$

where x is the distance from the mean of the item scores beyond which lies q proportion of the cases, σ_1 is the sigma of the criterion, and σ_2 is the sigma of the item scores.



Solving for $\overline{X}$,

$$\overline{X} = \frac{x\sigma_1}{r\sigma_2}$$

When the sigma of the item scores is 1, and the proportion of passes or failures as the case may be is q, x may be found from the *K-W* table and the formula becomes:

$$\overline{X} = \frac{x\sigma_1}{r}$$

Underlying the use of this formula is the assumption that normal correlation exists between the two variables.

The formula given above is the same as that published by Dr. Thorndike in "The Measurement of Intelligence," but differs in the principles underlying its derivation.

Dr. Thorndike shows that if the found validity of an item is less than perfect only because of a sampling error, then the persons scoring above the given scale point would *normally* pass the item and those scoring below the point would normally fail. The location of the point on the scale may be found by this formula.

Since he assumes the differentiation above and below this point to be perfect, if a normal distribution were selected such that fifty per cent would pass and fifty per cent fail, the mean of this distribution would have to coincide with this scale point. The fact that the formula has been derived in these two ways would seem to support the theory that while an item may not be regarded as a unit of measurment having a particular difficulty, it may be said to be part of a unit having that difficulty.

The percentage of passes, or functions of it, have for some time been regarded as indices of the difficulty of an item. Substituting the formula for biserial r (Kelley, p. 248) in the formula for difficulty, and assuming the sigma of the criterion to equal 1, we have

$$\frac{\frac{x}{pq(M-m)}}{z} = \frac{xz}{pq(M-m)}$$

$\frac{xz}{pq}$ is a function of the percentage of passes. $\frac{1}{(M-m)}$ is not only a function of the percentage of passes, but a function of the validity of the item. It will be seen therefore that if a series of items were arranged according to percentage of passes and according to the difficulty formula given here, there need not be very close correspondence between the two methods. Thurstone[1] has recognized the necessity for taking the validity of an item into consideration in determining its difficulty. The arrangement of items by the difficulty formula is probably to be preferred over that by percentage of passes because of the refinement of theory underlying it.

It is sometimes necessary to analyze an item on the basis of some group whose midpoint and variability differ from the midpoint and variability of the group for which a test is to be designed. Letting M be the mean and σ the sigma of the group analyzed, we have, by the formula $M + \frac{x\sigma}{r}$, the difficulty of the item in terms of the scale of the

[1] Thurstone, L. L.: Equally Often Noticed Differences. *Journal of Educational Psychology*, May, 1927.

criterion, upon which both groups are measured. This value of difficulty will be constant for any group on the scale, no matter what its midpoint or variability, providing it is a normal distribution. It is a much more complicated matter to calculate what the percentage of passes in a particular group will be, knowing the percentage of passes in some other group.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY JAMES E. MENDENHALL

Research Associate, Institute of Educational Research, Teachers College
Columbia University

Intelligence Testing

On Comparing IQ's at Different Age Levels on the Same Scale. Walter F. Dearborn and Howard H. Long. Journal of Educational Research, Nov., 1928, 265–274. The authors show that Stanford Binet IQ's must vary from year to year because of the different variabilities (SD's) for succeeding ages.

The Mental Age Concept and the Standardization of Group Tests. Godfrey H. Thomson. Psychological Review, Sept., 1928, 398–413. The author presents evidence to show that Thurstone's two definitions of mental age are practically identical; that is, for practical purposes predicting mental age from score is equivalent to predicting score from chronological age (as is done in calculating norms).

The Peg Board (Wallin) as a Means of Analyzing Form Perception and Motor Control in Young Children. Bird T. Baldwin and Beth L. Wellman. Journal of Genetic Psychology, Sept., 1928, 389–414. Four hundred ninety-eight series of tests were given by four examiners to two hundred sixty-nine children between the ages of two and six years. The results indicated (1) increased scores with increased ages; (2) no sex differences; (3) orderliness only in easy tasks; (4) greatest improvement on first repeated performance; (5) perseveration at early ages; and (6) correlations between first and second tests were low (−.05 to +.51).

The Relation of Enuresis to Intelligence, to Conduct and Personality Problems and to Other Factors. The Psychological Clinic, Sept.-Oct., 1928, 119–127. "As measured by its presence or absence among three thousand cases, enuresis showed a *small* negative correlation with intelligence quotient . . . There was a low positive correlation with both the number of conduct problems and the number of personality problems, but no significant correlation with underweight condition or with discord between parents."

A Preliminary Study of Intelligence as a Function of Birth Order. Harold E. Jones and H. H. Hsiao. Journal of Genetic Psychology, Sept., 1928, 428–433. A summary of some recent investigations.

A Report on the Examination of One Hundred 6B Children in Philadelphia Schools. Mabel R. Farson. The Psychological Clinic, Sept.-Oct., 1928, 128–152. A report of individual and group mental test results.

Why Are More Boys than Girls Retarded in School? (II) Ethel L. Cornell. The Elementary School Journal, Nov., 1928, 213–226. The girls surpass boys in "linguistic" tests (intelligence and reading) and ability to orient themselves to their surroundings, in stability and docility (personality ratings).

An Analytic Study of One Class in High School. Ann Eleanora Biddle. The Psychological Clinic, Sept.-Oct., 1928, 97–117. A study of two hundred five

entering ninth grade girls in a battery of performance, other mental tests, personality tests, in height and weight. The Binet memory span proved especially valid in predicting failure scholastically.

The Practical Application of the Student's Ability Index. Howard J. Banker, Journal of Educational Research, Nov., 1928, 282–289.

Measurement of Achievement

The Correlation between Measures of Mental Ability and Measures of Achievement in Chemistry. S. R. Powers. School Science and Mathematics, Dec., 1928, 981–986. The author reports the following correlations: (*N*, about 200) Powers Chemistry and (1) Terman Group IQ, .31, (2) Cleveland Intelligence Test, .46, (3) Otis Group IB, .67; (*N*, 50) Powers and (4) Otis Scores, .40, (5) Gerry Chemistry, .81, (6) Teacher's Marks, .53; Regent's and (7) Powers, .35, (8) Terman, .44.

A Study of a Modified Form of the Multiple-response Test. Francis D. Curtis and Gerald G. Woods. Journal of Educational Research, Oct., 1928, 211–219. The cases were two hundred six pupils in general science and biology, Grades VII to IX. In the modified form, the pupil either chooses the correct alternative or, if the correct alternative is not present in the test item, writes in the correct answer. This form was found more difficult for pupils, more stimulative to thinking, greater in diagnostic value than the conventional multiple choice test. The two forms were equivalent in reliability and validity.

Guessing in True-false Tests. Charles A. Leker. The School Review, Dec., 1928, 768–770. The study suggests that students exhibit tendencies from test to test of marking consistently a certain proportion of the test items "true" when guessing.

The Prognostic Value of General Language. Walter Kaulfers. School and Society, Nov. 24, 1928, 662–664. This report based upon one hundred eighty-six junior high school pupils presents a correlation between grades in general language and grades in French, Latin, and Spanish of .442. This coefficient is no higher than that between these languages and ordinary intelligence quotients.

A Study in Latin Prognosis. Jacob S. Orleans and Michael Solomon. The Classical Journal, Dec., 1928, 174–190. A total of three hundred fifty pupils were tested. Correlations between prognosis test and achievement tests ranged from .74 to .81.

The Textbook and Achievement in Latin. Jacob S. Orleans and Harold G. Thompson. School and Society, Nov. 3, 1928, 549–550. A testing of twenty-five hundred high school students revealed differences in achievement associated with the three texts used.

Do Algebra Students Need Remedial Arithmetic? George A. Boyce. School Science and Mathematics, Dec., 1928, 946–950. With a small group of high school students, correlations between different arithmetical processes and algebra scores ranged from .08 to .39.

Psychology of Learning and School Subjects

A Comparison of Two Methods of Learning. R. D. Russell. Journal of Educational Research, Oct., 1928, 235–238. A total of six hundred ninety pupils in the fifth, seventh, and ninth grades were tested. Results on objective tests showed that the teacher's oral reading of a passage was superior to the pupils' silent read-

ing of it themselves in the fifth grade; that the two methods were practically equal in effectiveness in the seventh grade; and that in the ninth learning by pupils' (silent) reading was slightly superior.

The Retention of American History in the Junior High School. Fowler D. Brooks and S. Janet Bassett. Journal of Educational Research, Oct., 1928, 195–202. "The present report shows the amount of American History actually retained by four hundred ninety-five students in Grades VII and VIII of several public schools in Baltimore from four to sixteen months after the semester in which it was studied. The results indicate that pupils forget twenty-three per cent of first or second-year junior high-school American History in a year's time."

Unfamiliar Situations as a Difficulty in Solving Arithmetic Problems. Carleton Washburne and Mable Vogel Morphett. Journal of Educational Research, Oct., 1928, 220–224. Eight pairs of problems, fifth grade level, were administered to four hundred forty-one children. The processes involved were column addition, subtraction with carrying, simple multiplication, short division, compound multiplication, and long division. According to *correct process*, the familiar showed a higher score in five of the eight problems, an average of five per cent in superiority; in *accuracy*, six of the eight problems favored the familiar, with an average superiority of six per cent.

The Vocabulary Load of Certain Secondary School Mathematics Textbooks. H. H. Remmers and A. Grant. Journal of Educational Research, Oct., 1928, 203–210. The authors present tables showing the vocabulary burden of twelve geometry and algebra books.

The Special Vocabulary of Civics. Orlando W. Stephenson. Journal of Educational Research, Nov., 1928, 297–304. A careful analysis of ten civics textbooks served as a basis for the vocabulary list.

Allies of the Ten Thousand. H. G. Good. Educational Research Bulletin, Nov. 28, 1928, 361–364. Using an analysis of literary, psychologic, and philosophic books, the author presents a list of words not occurring in the Teacher's Word Book. He makes a plea that the word book should be a guide, not a final source, for words.

Phonetic Spelling Scales. J. E. W. Wallin and Charlesanna B. Coles. Journal of Genetic Psychology, 434–470. The sigma scale was used to construct four equivalent sets of spelling scales for words differing in difficulty by .1, .2, and .3 sigma for each grade, II to VIII. The scales are included in the article.

Methology in Biology. Valeda G. Norris and George W. Hunter. School Science and Mathematics, Dec., 1928, 987–994. In a group of nine hundred sixty high school students, questionnaires show that about twenty per cent prefer "experiment or laboratory work done yourself;" twenty-five per cent, "experiments performed by teacher;" thirty per cent (girls) and ten per cent (boys), "field or museum trips;" ten per cent (girls) and fifteen per cent (boys), "class discussion on assigned lessons;" one per cent, "reference or outside reading;" fifteen per cent (girls) and twenty-six per cent (boys), "combination of methods."

Sarcasm. Thomas H. Briggs. The School Review, Nov., 1928, 685–695. Questionnaires submitted to high school and college students showed a belief that sarcasm hindered them from doing their best work.

In What Grade Should the Study of Foreign Language Begin? Elfrieda Steindorff. The Elementary School Journal, Nov., 1928, 209–212. "So far as this particular

school is concerned, a foreign language may be begun in the seventh grade as successfully as in the eighth or the ninth grade."

An Experimental Study of a Visual Method in Latin Instruction Norman Fenton and Victor D. Hill. The School Review, Nov., 1928, 675–678. With small groups (*N*, 9 each), the group using stereoptican slides proved superior in achievement.

Statistical Methods

The Method of Absolute Judgment in Psychophysics. Ernest Glen Wever and Karl Edward Zener. Psychological Review, Nov., 1928, 466–493. A study in weight discrimination to compare the method of absolute judgment with that of relative judgment.

A Table for Predicting the Validity and Reliability Coefficients of a Test When Lengthened. Harold A. Edgerton and Herbert A. Toops. Journal of Educational Research, Oct., 1928, 225–234.

Child Psychology

General Review and Summary: Child Psychology. B. T. Baldwin. Psychological Bulletin, Nov., 1928, 629–697. "A review of the literature, Jan. 1, 1923 to Mar. 31, 1928."

Corporal Punishment and Suspension. H. H. Davis. School and Society, Nov. 17, 1928, 632. The study of records in St. Louis shows a marked decrease per thousand pupils in corporal punishment from 141.1 in 1881 to 1.7 in 1925; in suspension, from 10.0 in 1881 to 2.0 in 1925. The first quarter of the school year exhibits the highest rates in both categories.

A Study of the Disciplinary Problems and the Method of Handling Them as Practiced in the Senior High School. John F. Pierce. The Teachers Journal and Abstract, Oct., 1928, 509–515. The investigation covers an analysis of magazine articles and of a questionnaire submitted to one hundred fifty high school principals.

Behavior Problems. William H. Kilpatrick. Childhood Education, Nov., 1928, 119–126. A psychologic and philosophic discussion of new methods for solving the maladjustments of children.

Fear. Harold Ellis Jones and Mary Cover Jones. Childhood Education, Nov., 1928, 136–143. A study of the reactions of young children, older children, and college students to a snake. "Children up to the age of two showed no fear of a snake; by three or three and a half, caution reactions were common . . . Definite fear behavior occurred more often after the age of four years, and was more pronounced in adults than in children."

Reading Disabilities in Relation to Maladjustment. Phyllis Blanchard. Mental Hygiene, Oct., 1928, 772–788. Four cases are reported in fairly complete detail.

Educational Psychology

Effect of Rest Periods upon Production. George H. Shepard. The Personnel Journal, Oct., 1928, 186–202. The results indicate that college students (men) on light-heavy muscular work cannot give maximum output unless rests of 16.7 per cent of the time during working hours are allowed.

Practice, Fatigue and Oscillation. A Study of Work at High Pressure. J. C. Flugel. The British Journal of Psychology Monograph Supplements, Vol.

XIII, 1928. Forty-six girls, ages nine years, five months to thirteen years, four months,were practiced in single column additions for forty-six days, twenty minutes per day. The author presents important data concerning fluctuations due to differences in ability, etc.

Eye Movement Duration, Pause Duration and Reading Time. Miles A. Tinker. Psychological Review, Sept., 1928, 385–397. Previous studies were analyzed. New data regarding the reading of prose, algebra, chemistry, and formulas are presented.

The Fundamental Error of the Concept of Purpose and the Trial and Error Fallacy. Zing Yang Kuo. Psychological Review, Sept., 1928, 414–433.

Administration and Supervision

Educational Magazines Read by Five Hundred Elementary-school Principals and Classroom Teachers. E. Lowell Kelley and Frederick L. Whitney. The Elementary School Journal, Nov., 1928, 176–180.

Labor Saving Devices Used in Office Administration in Secondary Schools. W. C. Reavis and Robert Woellner. The School Review, Dec., 1928, 736–744. Based on the tabulated results of a questionnaire submitted to five hundred twenty-two principals.

Office Hours of Secondary School Principals. W. C. Reavis and Robert Woellner. The School Review, Nov., 1928, 656–664.

Effect of the Summer Vacation on the Achievement of Pupils in the Fourth, Fifth and Sixth Grades. Elizabeth Bruene. Journal of Educational Research, Nov., 1928, 309–314. A study of achievement of twenty-six fifth and twenty-eight sixth grade pupils showed a gain over the vacation period in reading, and in nature study and science; a loss in arithmetic fundamentals and reasoning, in history and literature, in language usage, in spelling, and in total score.

How Much Time Is Required in the Fall for Pupils of the Elementary School to Reach Again the Spring Level of Achievement. M. J. Nelson. Journal of Educational Research, Nov., 1928, 305–308. "Grade VIIB (*N*, 49) had regained the spring level of achievement in *speed* in solving arithmetic problems involving only the fundamental processes, at the end of four weeks following the opening of school;" in problems solved *correctly*, at the end of six weeks. Grade VA required over six weeks to reach the spring level in arithmetic.

An Experiment of the Efficiency of Instruction in Large and Small Classes in Plane Geometry. Leonard D. Haertter. Educational Administration and Supervision, Nov., 1928, 580–590. The author concludes that this subject can be learned as well in a class of forty as in a class of twenty students.

A Study of the Relation between Teaching Efficiency and Amount of College Credit Earned While in Service. Harriet M. Barthelmess and Philip A. Boyer. Educational Administration and Supervision, Nov., 1928, 521–535. A total of five thousand two elementary and twelve hundred twenty junior high school teachers filled out a questionnaire, ratings on teacher efficiency being supplied by the principals. These data gave the following correlations: experience and credits, .228 (elementary) and .346 (junior high); experience and efficiency, .272 and .355; credits and efficiency, .16 and .189.

Educational and Vocational Guidance

Intelligence and Scholarship of Occupational Groups. L. D. Hartson. The Personnel Journal, Dec., 1928, 281–285. In psychological tests, the college teaching group (women and men) was highest, with physical education group lowest. In scholarship, the college teaching group was again highest, with business (men) and physical education (women) lowest.

Vocational Stability of Oberlin Alumni. L. D. Hartson. The Personnel Journal, Oct., 1928, 176–185. Returns from twenty-eight hundred sixty alumni reveal (separately for sexes) the occupational distributions by decade of graduation, by the time elapsed after graduation before entering final occupation, and by the first occupation.

Intelligence Test Scores and Choice of Major Field. K. C. Garrison. School and Society, Nov. 17, 1928, 630–632. Data on three hundred thirty-one students in southern teachers colleges reveal the highest Otis SA scores for (1) the *psychology*, secondary education, and religious majors; for (2) graduate students; and for (3) ages eighteen and twenty-three to twenty-six.

Predicting Success in First Semester College Social Science. M. Eustace Broom and Guy F. Fox. School and Society, Dec. 1, 1928, 689–690. With teachers college students (*N*, 259), average grade points per semester unit of credit in social science correlated with (1) average grade points per unit of high school social science, .207; (2) number of units in high school, .112; (3) Thorndike total scale, .230; (4) Thorndike reading decile rank, .266; and (5) Thorndike trade (information) decile rank, .089.

Parental Occupations and Vocational Choices of Private-school Students. Robert D. Cole. The Journal of Educational Sociology, Dec., 1927, 205–216.

How High School Pupils Spend Their Time. Inez M. Cook and T. V. Goodrich. The School Review, Dec., 1928, 771–778. The returns from questionnaires submitted to four hundred forty-one girls and to three hundred eighteen boys reveal distributions of time spent in various activities.

Measuring the Interest of Teachers in Their Work. Harry Dexter Kitson. Teachers College Record, Oct., 1928, 28–33. A ten step scale was devised for measuring degree of interest by self-rating methods.

A Test for Drivers. William Forster. The Personnel Journal, Oct., 1928, 161–171. Tests of intelligence, reaction time (special apparatus), and span of sensory perception were administered to one thousand individuals.

Predicting Success of Y. M. C. A. Secretaries. Harvey E. Becknell. The Personnel Journal, Oct., 1928, 172–175. By Yule's "significance difference" formula, the records of one hundred seventy-two general secretaries, rated as successful or as unsuccessful, were appraised for significant items of information.

The Problem Employee. V. V. Anderson. The Personnel Journal, Oct., 1928, 203–225. About one hundred twenty out of twelve hundred employees were studied by the psychiatrist in social history, job behavior, physical examination, and mental examination. Fourteen case histories are presented.

Study Habits and Methods

An Experimental Investigation of the Relative Effectiveness of Two Plans of Supervised Study. Harl R. Douglass. Journal of Educational Research, Oct.,

1928, 239–245. Ten pairs of classes (equated by Otis, CA, and initial achievement) were used as subjects. Study-recite sequence proved slightly superior for history, social studies, and literature; the recite-study, for mathematics and general science.

Diagnosis as a Basis for the Direction of Study. Maxie N. Woodring and Cecile White Flemming. Teachers College Record, Oct., 1928, 46–64, Nov., 1928, 134–147. In attacking this problem, records were secured of study activities, factors conditioning study: ability to learn, ability to read, achievement in the specific subjects, physical condition, home study background, school study background, schedule of work in and out of class, methods of work, associates, attitude of pupil toward work, personality of teacher, assignments, etc.

Educational Sociology

Determining the Results of Education. E. George Payne. The Journal of Educational Sociology, Apr., 1928, 468–476. A controlled experiment of changes in accident rates and in physiological conditions of pupils due to education.

The Socio-economic Status of Three School Populations. Verner Martin Sims. The Journal of Educational Sociology, Oct., 1928, 83–91. The study presents the differences between city, town, and country groups of children in their performances on the author's score card.

Miscellaneous

Curriculum Research. Edward L. Thorndike. School and Society, Nov. 10, 1928, 269–576. The author points out the bearing of certain significant scientific studies in education upon various aspects of the curriculum.

Measuring Musical Talent. Seashore Tests as Administrative Aids. Hazel M. Stanton. The Personnel Journal, Dec., 1928, 286–292. A study of various uses of the tests for admissions, eliminations, relative permanence of students, etc.

A Critical Study in the Objective Measurement of Character. F. J. Brown and Marguerite Shelmadine. Journal of Educational Research, Nov., 1928, 290–296. Two groups, one hundred twenty-four high school and one hundred eighty-six junior high school pupils, were tested in knowledge of honest responses and on paper-pencil honesty tests.

Age, Height, Weight and Pubescence Standards for the Athletic Handicapping of Girls. Mary Delaney. American Physical Education Review, Oct., 1928, 507–509. The athletic records of eight hundred thirty-three girls between the ages of ten and sixteen and one-half years were collected from the seventh and eighth grades. Correlation techniques were applied to these data.

Do College Students Know and Understand the Lord's Prayer? Lester R. Wheeler. School and Society, Dec. 8, 1928, 735–736. Of five hundred forty teachers college students, this study shows that a much larger percentage can repeat the Lord's Prayer (ninety-five to ninety-seven per cent), than can understand the thought units (forty-four to eighty-two per cent).

Statistical Analyses of the Members and Associates of the American Psychological Association, Inc. in 1928. Samuel W. Fernberger. Psychological Review, Nov., 1928, 447–465.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

Conduct Made-to-order

The Nature of Conduct, by Percival M. Symonds. New York: The Macmillan Co., 1928. Pp. IX + 346.

Conduct comprises the habits and skills with which human beings meet life. If all of conduct be taken together, it becomes character. All of the fine enthusiasm which educators are showing for character education must be focussed by a realization that the ideals they seek are to be attained in terms of specific situations.

Since "careful analysis reveals nothing in the body besides the nervous system which could act as a control of activity, all the intricate problems of choice, purpose, reasoning, problem solving, and the like must eventually be explained by neural paths and synaptic resistances." Appropriately, therefore, the book offers an excellent brief summary of the structure and operation of the nervous system.

In the analysis of the common elements which make certain behavior situations seem "alike" and which have led some to speak of moral traits, the author offers a useful symbol in his term "confact." Just as concepts represent the verbal response to subtle factors in several situations, so confacts represent the conduct responses to those subtle factors.

Stimulating challenge to many character education procedures is offered by such dogmas as: the formation of conduct begins with separate skills and habits, not proceeding from the general down to particulars; words are far from magical controls in conduct, being useful only in those rare circumstances in which learning has connected the idea with the act; reasoning about conduct seldom influences it; pupils cannot be expected, in general, to solve conduct problems by analysis of consequences; inner attitudes and feelings about conduct are not the springs of action, but mere ineffective concomitants; conduct for most people remains unorganized, generalization being far less extensive than is popularly supposed. While it is recognized

that new occasions will demand new duties the author's position, in general, is that the desired behaviors in life must be listed, defined, practiced, and so developed into habits. He offers the interesting, incidental suggestion that a new series of symbols be developed, arousing the same emotional and motor loyalty which the symbols of religion have had, but connected in logical fashion with the conduct responses which our society finds by scientific analysis to be good. Science at the work-table is to fashion a made-to-order religion.

GOODWIN B. WATSON.

Teachers College, New York.

THE MIND BEHIND THE TONGUE

The Psychology of Language by Walter B. Pillsbury and Clarence L. Meader, New York: D. Appleton and Co., 1928. Pp. VII + 306.

If the centipede really lost his ability to walk when he stopped to consider the order of precedence of his legs, then the layman best be warned against reading the chapters in the "Psychology of Language" which deal with the motor processes involved in speech. Dismayed by the complication and delicacy of the adjustments involved, he might conceivably lose his ability to speak. A series of carefully made measurements indicates, for example, that in making successively the two sounds, of *i* in *pin* and *a* in *bake*, the teeth change their separation by $^{5}/_{1000}$ of an inch, and the other essential adjustments are on an equally minute scale.

The authors have attempted to bring together in their book all of the materials which have to do with the subject of language in its various aspects and to give to them a psychological interpretation. Their own statement in the preface is to the effect that the book "attempts to give a full description of those phases of psychology which have a bearing on the general phenomena of language and to treat the accepted laws of language from a psychological point of view." Contemporary thought in the field being what it is, the reconciliation of conflicting points of view occupies quite as much of the text as does interpretation of laws concerning which substantial agreement exists.

Successive chapters of the book deal with the physiology of the sense and motor organs involved in speaking, listening and reading, and with the psychology of perception and response; with the gram-

matical analysis of speech units; with the history of the origin of language; with a philosophical consideration of the meaning of meaning. Included in each of the chapters is some discussion by the authors of the material cited from the studies which have been made of these various topics. The discussion lacks the guidance of a fixed point of view consistently adhered to. If the authors believe that speech is to be considered simply as one highly specialized, complicated, and interesting reaction among the total possible responses of which the articulate human is capable in the presence of stimulating situations including, among other factors, the written page and the speech of other persons, and that these reactions are subject fundamentally to the same explanations as are all others—they fail to insist upon this point of view. The upshot of the discussion is not made clear. Summaries, particularly important in a work which is essentially a compilation of contributions from a large variety of sources, are lacking. Chapters end abruptly, leaving the reader with the feeling that he has arrived nowhere in particular. The final chapter ends with the cryptic statement, "All is one," a declaration of monistic faith which might be illuminating in respect to the language arts if the reader had been given a definite notion of what the magically explanatory "one" is.

The authors have a certain optimism about the language arts, illustrated by the statement that "The first essential of speech is something to say. After speech has been learned, talking is simply thinking in words with the intention of expressing one's self. The first essential is that words should be suggested. When once suggested the speaking begins." One cannot help believing that a great deal of speech and writing goes on in the absence of anything particular to say, and to conclude that all speech is for the purpose of self-expression is to subscribe to a teleology which the facts do not warrant. All sorts of intentions other than the intention to speak may easily set off the speech mechanism, and no small item of the total speech production can reasonably be attributed to stimulating situations other than intention.

The form of the book is marred by the monotony of sentence structure to which one would expect that persons expert in the psychology of speech would have been sensitive. Chapters IX and X, by rough count, run about four sentences to the page beginning "It is" and "There are." The impersonality of statement attained by the use of this inverted form of sentence adds to the impression of lack of

definiteness resulting from the absence of chapter summaries referred to above.

As a handbook of data on the psychology of the language arts, the book will be useful to persons interested in that field. Using the quite adequate index, the student can inform himself on any particular aspect of the subject in terms of what has been said and written by other persons without the labor of searching source materials.

LESTER WILSON.
International Institute, Teachers College, New York.

A CODE OF ETHICS

Natural Conduct by Edwin Bingham Copeland. Stanford University Press, 1928. Pp. VII + 262. $3.50.

"Natural Conduct" is a book with a purpose. That purpose is to present what the author thinks is a code of ethics based on the teachings of modern biology. The basic assumption, of course, is that mankind can best be improved not by eugenics or collective education but by the ethical education of the individual. The intellectual background is an interest in problems of ethics by way of "family inheritance" which was "intensified and directed in the undergraduate days by Doctor Jordan's teaching of organic evolution." The mechanical make-up of the book is excellent; the style is clear and correct.

The approach is intellectualistic. The thesis is stated in terms of Aristotle's axiom to the effect that it is the goodness of things which preserves them. It may be said in passing that the author accepts the logical implication of this statement which readers of James Harvey Robinson will find hard to believe; namely, that "there is in general a complete agreement between the rules of conduct which are accepted on faith and the rules of conduct which can be deduced by the processes of reason." The infallibility of nature is, in a way which reminds one of the writings of G. Stanley Hall, tacitly assumed throughout the consideration of all problems including, "The Family," "Duty and Happiness," "Ideals," "Proper Conduct," "The State," "The School." Concepts from faculty psychology, rationalizations about our feelings and emotions, hearsay stories about the behavior of wolf-packs—all these can be found in this book.

It may be as the author assures us that "as we strengthen ourselves in wisdom, the control of our conduct will be safe, our conduct will be

worthy, and our days will be completely happy." His code of ethics if followed may lead to safety, conformity, complacency. The education advocated may be ideal for a contemplative life. But man is more than his intellect; mind is more than intelligence. "A pasteurized and sanitized society is not," in the words of Todd, "necessarily a progressive or dynamic one." That is why the reaction of the reviewer to this book can be best expressed in the manner of the reaction of William James after a visit to a Chatauqua: Ouf! what a relief! This order is too tame; this goodness too uninspiring. Give me Ellis, Russell, Dewey, Haldane who, with Roscoe Pound, appreciate the wholeness of life and with him insist that "we must build rather than merely approve; we must create rather than merely order and systematize and logically reconcile details."

H. MELTZER.
Psychiatric Clinic, St. Louis.

PROBLEMS OF YOUTH

The Psychology of the Adolescent, by Leta S. Hollingworth. New York: D. Appleton and Co., 1928. Pp. 259.

In recent years the point of view has almost universally been adopted that the phenomena of adolescence are due primarily to the facing of new problems of life which present themselves at this time, rather than to any sudden blossoming forth of new traits of the individual. A modern study of the adolescent thus draws as much from sociology and anthropology as from psychological theory. The problems which the adolescent meets are the precipitating events, not changes in the adolescent himself.

Dr. Hollingworth's book might have been subtitled "The Persistent Problems of Youth." Each of these problems, such as that of getting away from home and family—very aptly termed Psychological Weaning—is developed from its status and solution in primitive tribes to the complexities of modern life. Although the attack is therefor principally historical and practical, the title of the book as a Psychology is not too much in error. The solving of the adolescent's problems is treated from a standpoint of habit formation, without the necessity of unique hypotheses. The relationship between level of intelligence and the occurrence of the problems is also adequately taken into account.

The organization of the book is very adaptable. No graphs, references or exercises disturb the continuity of the text for the non-

scholastic reader. For those who wish a fuller study, or for class use, these devices are very completely provided in three appendices.

Dr. Hollingworth has written principally for the parent or teacher. The older adolescent of college years who is himself in the midst of the adjustments considered will, however, find the descriptions intriguingly applicable, and will perhaps profit most. LAURANCE F. SHAFFER.

Carnegie Institute of Technology.

A MEATY BOOK FOR THE WORKER IN SECONDARY EDUCATION

The High School, by Walter S. Monroe and Oscar F. Weber. Garden City, New York; Doubleday, Doran and Co., Inc., 1928. Pp. VIII + 511.

Here is a book which will be a boon to many a high school principal and his staff. While the authors wrote it primarily for students of secondary education (and it will serve them well), it is admirably adapted for a project by those secondary school staffs whose teachers' meetings consist of something more than a list of announcements. This reviewer read it with sustained interest and admiration. It is a book which present and prospective secondary school teachers can study with pride in their chosen work, with a growing consciousness of secondary school teaching as a profession, and with increasing understanding of the meaning and problems of secondary education in the United States today.

Leaving the psychology of secondary school subjects and methods of teaching to other texts, this book considers what is usually meant by principles of secondary education. The first one hundred twenty-one pages are worth the price of the text. They deal with the development of secondary education in the United States; contemporary secondary education in England, France, and Germany; and the present downward and upward extension of secondary education in the United States. The treatment is not long but thorough while excellent bibliographies suggest more intensive studies for the student. Throughout the text the bibliographies are good.

Those schools which are undertaking the reconstruction of their curricula will find this text a good point of departure. The following point of view is found on page 224: "In general the high school teacher is not expected to engage extensively in curriculum construction, but it is highly desirable that he maintain the right attitude toward pro-

posals for changes and be able to evaluate them wisely." To this end, the recent history of the aims and practices of each of the major subjects of secondary schools is discussed. We predict that most of the teachers in any department will find the treatment of their specialty stimulating. Most language teachers will be interested in the discussion of the General Language Tryout Course treated on p. 275. One regrets the authors did not point out in the chapter on mathematics that while various national committees were attempting to determine more or less empirically the requirements for courses in mathematics, Thorndike was pointing out a more promising procedure in his "Psychology of Algebra;" or, in the discussion of chemistry, that Powers' findings concerning the traditional course were not mentioned. Each of these is particularly important in its own field. Their omission does not constitute a basic fault in the text. The authors were conscious of the complex problems of secondary education today. Those who use the text should become conscious of them also, become familiar with most of the authorities in the field, and be stimulated to helpful thinking since the purpose of raising the important questions (to which final blanket answers cannot be given) has been realized. What principal will not be happy to refer his teachers to the chapter on the administration of the high school? With due consideration for most school teachers' salaries, it can be said earnestly that this is a book every secondary school teacher can little afford not to own.

J. H. Coleman.
Graduate Student, Teachers College, N. Y.